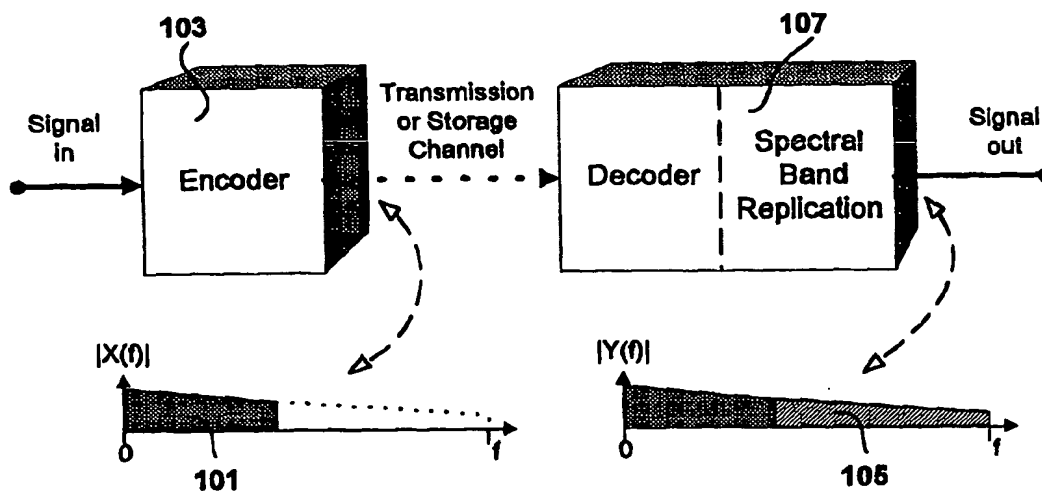


PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : H04B 1/66		A2	(11) International Publication Number: WO 98/57436
			(43) International Publication Date: 17 December 1998 (17.12.98)
(21) International Application Number: PCT/IB98/00893		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 9 June 1998 (09.06.98)			
(30) Priority Data: 9702213-1 10 June 1997 (10.06.97) SE 9704634-6 12 December 1997 (12.12.97) SE 9800268-6 30 January 1998 (30.01.98) SE			
(71)(72) Applicant and Inventor: LILJERYD, Lars, Gustaf [SE/SE]; Vintervagen 19, S-171 34 Solna (SE).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): EKSTRAND, Per, Rune, Albin [SE/SE]; Renstiernas Gata 29, S-116 31 Stockholm (SE). HENN, Lars, Fredrik [CH/SE]; Ritarvagen 14, S-168 31 Bromma (SE). KJÖRLING, Hans, Magnus, Kristofer [SE/SE]; Vindhemsgatan 19C, S-752 27 Uppsala (SE).			
(74) Common Representative: LILJERYD, Lars, Gustaf; Vintervagen 19, S-171 34 Solna (SE).			
		Published Without international search report and to be republished upon receipt of that report.	

(54) Title: SOURCE CODING ENHANCEMENT USING SPECTRAL-BAND REPLICATION



(57) Abstract

The present invention proposes a new method and apparatus for the enhancement of source coding systems. The invention employs bandwidth reduction (101) prior to or in the encoder (103), followed by spectral-band replication (105) at the decoder (107). This is accomplished by the use of new transposition methods, in combination with spectral envelope adjustments. Reduced bitrate at a given perceptual quality or an improved perceptual quality at a given bitrate is offered. The invention is preferably integrated in a hardware or software codec, but can also be implemented as a separate processor in combination with a codec. The invention offers substantial improvements practically independent of codec type and technological progress.



THIS PAGE BLANK (USPTO)

(19) 日本国特許庁 (J P) (12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2001-521648

(P2001-521648A)

(43) 公表日 平成13年11月6日 (2001.11.6)

(51) Int. Cl. ⁷	識別記号	F I	チャート (参考)
G 1 0 L 19/02		H 0 3 M 7/30	A
	19/00	H 0 4 B 1/66	
H 0 3 M 7/30		G 1 0 L 7/04	G
H 0 4 B 1/66			M

審査請求 有 予備審査請求 有 (金 79 頁)

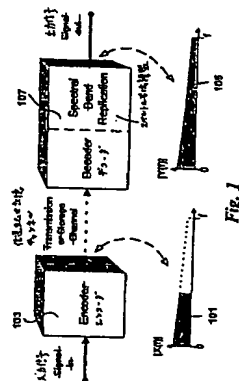
(21) 出願番号	特願平11-501962	(71) 出願人	コーディング テクノロジーズ スウェーデン アクチボラゲット
(86) (22) 出願日	平成10年6月9日 (1998.6.9)		スウェーデン国 ストックホルム スベアベゲン 119
(85) 優先出願番号	P C T / I B 9 8 / 0 0 8 9 3	(72) 発明者	リルイェリド, ラルス, グスタフ
(87) 国際公開番号	W O 9 8 / 5 7 4 3 6		スウェーデン国 エス-71 34 ソルナ, ピンデルバゲン 19
(87) 国際公開日	平成10年12月17日 (1998.12.17)	(72) 発明者	エクストランド, ベル, ルネ, アルビン
(31) 優先権主張番号	9 7 0 2 2 1 3 - 1		スウェーデン国 エス-116 31 ストックホルム, レンステイエルナス ガタ 29
(32) 優先日	平成9年6月10日 (1997.6.10)	(74) 代理人	井理士 浅村 皓 (特 3 名)
(31) 優先権主張番号	9 7 0 4 6 3 4 - 6		
(32) 優先日	平成9年12月12日 (1997.12.12)		
(33) 優先権主張国	スウェーデン (S E)		

(54) [発明の名称] スペクトル帯域複数をを用いた原始コーディングの強化

最終頁に続く

(57) [要約]

本発明は、原始コーディング装置を強化するための新しい方法と装置を示す。本発明は、エンコーダ (103) の前または中の帯域幅の削減 (101) と、その後のデコーダ (107) でのスペクトル帯域の複製 (105) を用いる。これは、新しい複製法とスペクトル帯域複数を組み合わせて行う。所定の知覚品質でビットレートを減らすか、または所定のビットレートで知覚品質を高める。本発明は好ましくはハードウェアまたはソフトウェアコーディングに組み込むが、別個のプロセッサとコーデックを組み合わせて実現してもよい。本発明は、コーデックの種類や技術的進歩とは独立に、実際に実質的な改善を与える。



[特許請求の範囲]

1. 原始コーディング装置の強化方法であって、前記原始コーディング装置は、記号または伝送の前に行う全ての操作を表すエンコーダと、記号または伝送の後に行う全ての操作を表すデコーダを含み、前記エンコーダで初期の信号の周波数帯域を複製して第1信号を形成し、前記デコーダで前記第1信号に複製を行って前記初期の信号の周波数帯域を複製して第2信号を形成し、前記第1と第2信号を結合して出力信号を形成し、これにより所定の知覚品質でビットレートを下げ、または所定のビットレートで知覚品質を高める、ことを特徴とする、原始コーディング装置の強化方法。
2. 前記第2信号の通過帯域は前記第1信号の通過帯域と重ならないまたは一部だけ重なるように設定することを特徴とする、請求項1に記載の原始コーディング装置の強化方法。
3. スペクトル包絡線調整は、前記第1信号を用いて、前記初期の信号の前記複製された周波数帯域のスペクトル包絡線の推定に基づいて行うことを特徴とする、請求項1-2に記載の原始コーディング装置の強化方法。
4. スペクトル包絡線調整は、前記初期の信号の前記複製された周波数帯域の、伝送された包絡線情報に基づいて行うことを特徴とする、請求項1-2に記載の原始コーディング装置の強化方法。
5. 前記スペクトル包絡線情報は、その利得が低レベルに設定された任意の数のサブバンドチャネル内のサブバンドサンプルとして伝送され、これにより標準化されたデコーダとの互換性が確保されることを特徴とする、請求項4に記載の原始コーディング装置の強化方法。
6. 前記包絡線情報を換算係数として伝送し、対応するサブバンドサンプルは伝送しないことを特徴とする、請求項4に記載の原始コーディング装置の強化方法。
7. 前記包絡線情報を換算係数として伝送し、前記対応するサブバンドサンプルをゼロまたは一定値に設定することにより、前記サブバンドサンプルのエン

トロピーを減らすことを特徴とする。請求項4に記載の原始コーディング装置の強化方法。

8. モノフォニックオクターブのときは、前記出力信号を、前記出力信号とそれを増強した信号をそれぞれ含む2つの信号に分割して疑似ステレオ信号を得ることを特徴とする。請求項1-7に記載の原始コーディング装置の強化方法。

9. 前記置換は、信号を、それぞれ周波数 $[f_1, \dots, f_N]$ を含む通過帯域を持つ N 個 ($N \geq 2$) の帯域フィルタの集合で透過して、 N 個の帯域信号を形成し、前記帯域信号の周波数を、周波数 $M[f_1, \dots, f_N]$ を含む領域にシフトし (ただし、 $M \neq 1$ は置換係数)。

前記シフトされた帯域信号を結合して置換された信号を形成する。ことを特徴とする。請求項1-7に記載の原始コーディング装置の強化方法。

10. 前記周波数シフトを上側帯域 (USB) 変調により得ることを特徴とする。請求項9に記載の原始コーディング装置の強化方法。

11. 係数 M で置換する方法であって、信号を、低帯域型の実数値または複素値のサブバンド信号を生成する性質の分解フィルタバンクまたは変換を用いて帯域透過し、

合成フィルタバンクまたは変換内で、前記分析フィルタバンクまたは変換の任意の数のチャネル k をチャネル Mk ($M \neq 1$) にバッチングし、

前記合成フィルタバンクまたは変換を用いて、置換された信号を形成する、ことを特徴とする。係数 M で置換する方法。

12. 前記フィルタバンクを最大10進化し、前記バッチングを次の関係により行い、

13. 前記分析フィルタバンクまたは変換のチャネル k からのサブバンド信号の位相を、合成チャネル Mk ($M \neq 1$) に関連するサブバンド信号の位相としてバッチングし、

前記分析フィルタバンクまたは変換の連続的なチャネル l からのサブバンド信号の振幅を、連続的な合成チャネル $l+S$ (S は整数 $\neq 1$) に関連するサブバンド信号の振幅としてバッチングする。

ことを特徴とする。請求項11-12に記載の係数 M で置換する方法。

14. 前記合成フィルタバンクまたは変換を用いる前に、前記チャネル k の前記サブバンド信号の位相に前記係数 M を掛けることを特徴とする。請求項11-13に記載の係数 M で置換する方法。

15. $M = K \pm 1$ (ただし、 K は整数 > 1) であることを特徴とする。請求項11-14に記載の係数 M で置換する方法。

16. 前記バッチングは前記置換係数 M の多重の値を用いることを特徴とする。請求項11-15に記載の係数 M で置換する方法。

17. 係数 M で置換する方法であって、インパルス応答

$$h(n) = K m(n) \exp \left[j \frac{\pi}{2L} (2L+1)n + \pi(n - \frac{M-1}{2}) + j(-1)^n \frac{\pi}{4M} \right].$$

ただし、 $k=0, 1, \dots, L-1$, K は定数、 $p_0(n)$ は長さ N の低域プロトタイプフィルタ、を持つ L 個のフィルタの並列バンクで信号を透過して、 L 個の複素値信号の集合を生成し、

係数 L/M を持つ前記 L 個の信号をダウンサンプリングして、 L 個の複素値サブバンド信号の集合を生成し、

前記複素値サブバンド信号の位相角に M を掛けて、サブバンド信号の新しい集合を生成し、

前記サブバンド信号の新しい集合の実数部を選択して、 L 個の実数値サブバンド信号の集合を生成し、

係数 L' を持つ前記実数値サブバンド信号の部分集合をアップサンプリングして、実数値信号の集合を生成し、

ただし、 $(-1)^{(M-1)kn}$ は訂正係数、 $v_k(n)$ はチャネル k のサブバンド信号、 $v_{Mk}(n)$ はチャネル Mk のサブバンド信号であり、これによりスペクトル反転サブバンド信号の補正が得られることを特徴とする。請求項11に記載の係数 M で置換する方法。

$$v_{Mk}(n) = (-1)^{(M-1)kn} v_k(n).$$

インパルス応答

$$A(n) = K' \rho(n) \cos \left[\frac{\pi}{2L'} (2k+1) \left(\rho - \frac{\rho' - 1}{2} \right) - (n) \frac{\pi}{4} \right].$$

ただし、 $k=0, 1, \dots, L'-1$ 、 K' は定数、 ρ, ρ' 、 $\rho(n)$ は長さ N' の低域プロトタイプフィルタ、を持つ L' 個のフィルタの並列バンクで前記乗数値信号を構成して、 L' 個の乗数値信号の集合を形成し、

前記 L' 個の乗数値信号を加算して乗換信号を生成する、

ことを特徴とする、係数 M で乗換する方法、

18. 前記位相角の前記掛け算と前記乗数値部の前記選択を計算するの、前記乗数値サブバンド信号を次式で置き、

$$Z_i(n) = R_k(n) + jI_k(n),$$

ただし、 $R_k(n)$ と $I_k(n)$ はそれぞれ $Z_k(n)$ の実数部と虚数部であり、前記乗数値サブバンド信号 $W_k(n)$ を次式で計算し、

$$W_k(n) = |Z_k(n)| \cos \left\{ M \arctan \left(\frac{I_k(n)}{R_k(n)} \right) \right\}.$$

ただし、 $|Z_k(n)| = \sqrt{R_k(n)^2 + I_k(n)^2}$ 、 M は正の整数の乗換係数であり、次の三角恒等式

$$\cos(M\alpha) = \cos^M(\alpha) - \binom{M}{2} \sin^2(\alpha) \cos^{M-2}(\alpha) + \binom{M}{4} \sin^4(\alpha) \cos^{M-4}(\alpha) - \dots,$$

ただし、 $\alpha = \arctan(I_k(n)/R_k(n))$ 、と次の関係

$$\cos(\alpha) = \frac{R_k(n)}{|Z_k(n)|} \text{ and } \sin(\alpha) = \frac{I_k(n)}{|Z_k(n)|}$$

を用い、これにより全ての三角法計算をなくして計算の複雑さを減らす、ことを特徴とする請求項 17 に記載の係数 M で乗換する方法、

19. ブロック毎に、前記乗数値サブバンド信号の隣接対の位相差により選ばれる情報を取り出し、

前記位相角に前記 M を掛けて前記新しいサブバンド信号の対を形成し、前記情報を与える条件によって前記新しいサブバンド信号の 1 つを打ち消すこ

とにより、偶数乗数値の乗換係数 M を用いるときにサブバンド信号の 180° 位相シフトを保持する、

ことを特徴とする、請求項 17 に記載の係数 M で乗換する方法、

20. 前記情報は次式の前記乗数値サブバンド信号 $Z_k(n)$ と $Z_{k+1}(n)$ の点乗積で与えられ、

$$Z_i(n) \cdot Z_{k+1}(n) = R_k(n) R_{k+1}(n) + j I_k(n) I_{k+1}(n),$$

ただし、 $R_i(n)$ と $I_i(n)$ はそれぞれ $Z_i(n)$ の実数部と虚数部 ($i = k, k+1$) であり、前記点乗積が負の場合は前記新しいサブバンド信号の 1 つを打ち消すことを特徴とする、請求項 19 に記載の係数 M で乗換する方法、

21. 第 1 信号を時間的に伸張または圧縮し、前記第 1 信号の任意の長さのセグメントを複写または圧縮し、次に前記第 1 信号をダウンサンプリングまたはアップサンプリングする、乗換方法であって、

前記第 1 信号に過渡検出を行い、

過渡検出の結果に従って、前記第 1 信号の一部を複写または圧縮するときに前記第 1 信号のどのセグメントを用いるかを決定し、

前記過渡検出の結果に従って前記信号セグメントの長さ L を調整し、

前記過渡検出の結果に従って各状態ベクトルに用いるサンプル数 N を調整し、

前記過渡検出の結果に従って状態ベクトル内のサンプル間の遅れ D を調整し、

前記過渡検出の結果に従って各状態ベクトル間のサンプル数 K を調整し、

前の同期点探索で見出した同期点に基づいて、前記第 1 信号の選択されたセグメント内の同期点を探索、

ことを特徴とする乗換方法、

22. いくつかのトランスフォーマーを相互接続して同期点情報を共有して、計算の複雑さを減らすことを特徴とする、請求項 21 に記載の乗換方法、

23. 前記トランスフォーマーを適当なフィルタバンクに接続し、前記各トランスフォーマーに与える信号を減波して、前記トランスフォーマーが処理中の前記信号の和である新しい信号の任意のスペクトル包絡線を得ることを特徴とする、請求項 21-22 に記載の乗換方法、

前記エンコーダで前記初めの信号のスペクトル包絡線情報を取り出して第2信号を形成する取出し手段と、

前記エンコーダで前記第1信号と第2信号を符号化する手段と、

前記デコーダで前記第1信号の周波数帯域を置換して第3信号を形成する置換手段と、

前記第2信号に基づいて、前記デコーダで前記第3信号のスペクトル包絡線を調整する調整手段と、

前記デコダで前記第1番号と前記調整された第3番号を結合して、所定の知覚品質でビットレートを下げ、または所定のビットレートで知覚品質を高める、結合手段。

である原始コーディングの強化装置。

27. 前記出力信号がモノフォニックオーディオのときに動作し、

第1遅延信号を形成するための、前記出力信号を遅らせる遅延手段および減衰させる減衰手段と、

第2遅延信号を形成するための、異なるパラメータを用いる、前記出力信号を遅らせる遅延手段および減衰させる減衰手段と、

と、前記出力と前記第 1 遅延信号を加算して左チャンネル出力信号を形成する手段

前記出力と前記第2遅延信号を加算して右チャネル出力信号を形成して、疑似ステレオフォーマット信号を得る手段、

を特徴とする、請求項26に記載の原始コーディングの強化装置。

28. 係数Mで置換する装置であつて、

信号を、低帯域型の実数値または複素数値サブバンド信号を生成する性質の分析フィルタバンクまたは変換により帯域濾波すること、

合成フィルタバンクまたは変換内で、前記分析フィルタバンクまたは変換の任意のチャネル数 k をチャネル Mk ($M \neq 1$) にパッチングする手段と、

前記合成フィルタバンクまたは変換により、置換された信号を形成すること、
を特徴とする係数Mで置換する装置。

29. $M=K \pm 1$ のとき (K は整数 > 1)、前記合成フィルタバンクまたは変換を用いる前に前記チャネル k のサブバンド信号の位相に M を掛けることを特徴とする。請求項 28 に記載の係数 M で置換する装置。

30. 係数 M で置換する装置であって、インパルス応答

$$h_0(n) = \sum_{k=0}^{L-1} p_k(n) \exp \left[j \frac{\pi}{2L} (2k+1) \left(n - \frac{N-1}{2} \right) + j(-1)^k \frac{\pi}{4M} \right],$$

ただし、 $k=0, 1, \dots, L-1$, $L-LK$ は定数、 $p_0(n)$ は長さ N の低域プロトタイプフィルタ、 M は置換係数、を持つ L 個のフィルタの並列バンクで信号を濾波して、 L 個の複素値信号の集合を生成する濾波手段と、係数 L/M を持つ前記 L 個の信号をダウンサンプリングして、 L 個の複素値のサブバンド信号の集合を生成する手段と、前記複素値サブバンド信号の位相角に M を掛けて、サブバンド信号の新しい集合を生成する手段と、前記サブバンド信号の新しい集合の実数部を選択して、 L 個の実数値サブバンド信号の集合を生成する手段と、係数 L' を持つ前記実数値サブバンド信号の部分集合をアップサンプリングして、実数値信号の集合を生成する手段と、インパルス応答

$$h_1(n) = K' p_0(n) \exp \left[j \frac{\pi}{2L'} (2k+1) \left(n - \frac{N'-1}{2} \right) + j(-1)^k \frac{\pi}{4} \right],$$

ただし、 $k=0, 1, \dots, L'-1$, $L'-1$, K' は定数、 $p'_0(n)$ は長さ N' の低域プロトタイプフィルタ、を持つ L' 個のフィルタの並列バンクで前記実数値信号を濾波して、 L' 個の複素値信号の集合を形成する濾波手段と、前記 L' 個の複素値信号を加算して置換信号を生成する手段、を特徴とする。係数 M で置換する装置。

31. 第 1 信号を時間的に伸縮または圧縮し、前記第 1 信号の任意の長さのセグメントを複写または複製し、次に前記第 1 信号をダウンサンプリングまたは

アップサンプリングする、置換装置であって、前記第 1 信号に過渡抽出を行う後出手段と、可能な過渡信号の位置を用いて、前記第 1 信号の一部を複写または複製すると共に前記第 1 信号のどのセグメントを用いるかを決定して、前記置換を得る手段と、

と、前記過渡抽出器からの出力に従って前記信号セグメントの長さ (L) を調整する調整手段と、前記過渡抽出器からの出力に従って各状態ベクトルに用いるサンプル数 (N) を調整する調整手段と、前記過渡抽出器からの出力に従って前記状態ベクトル内のサンプル間の遅れ (D) を調整する調整手段と、前記過渡抽出器からの出力に従って各状態ベクトル間のサンプル数 (K) を調整する調整手段と、前の同相点探索で見出した同相点に基づいて、前記第 1 信号の選択されたセグメント内の同相点を探索する探索手段、を特徴とする置換装置。

32. サブバンド信号に作用して、前記トランスフォーマーの多数の事例の間で同期情報を共用する手段と、前記サブバンド信号の部分集合を形成する手段と、前記各部分集合内で各チャネルの振幅調整を行う手段と、前記各部分集合から、前記トランスフォーマーの各事例への入力信号を形成する合成フィルタバンク手段と、前記トランスフォーマーによる前記入力信号の処理と、前記処理信号を加算することにより新しい信号を得て、任意のスペクトル包絡を得る加算手段、を特徴とする。請求項 31 に記載の置換装置。

【発明の詳細な説明】

スベクトル帯域複製を用いた原始コーデイングの強化

技術分野

原始コーデイング装置では、必要なビットレートや記憶容量を減らすためにデジタルデータを圧縮して伝送または記憶する。本発明は、スベクトル帯域複製(SBR)により原始コーデイング(source coding)を改善する新しい方法と装置に関するものである。同じ知覚品質を保持してビットレートを実質的に下げ、逆に所定のビットレートで知覚品質を高める。これは、エンコード側でスペクトル帯域を縮小し、デコード側で後のスベクトル帯域を複製することにより行う。本発明はスベクトル帯域での信号冗長度の新しい概念を活用する。

発明の背景

オーディオ原始コーデイング技術は2種類ある。すなわち、自然オーディオコーデイングと音声コーデイングである。自然オーディオコーデイングは中位のビットレートの音楽や言葉の信号に共通に用いられており、オーディオ帯域幅は一般に広い。音声コーデイングは基本的に音声の再生に限られるが、他方では非常に低いビットレートで用いることができる。ただしオーディオ帯域幅は狭い。広帯域音声は狭帯域音声に比べて主な主観的品質が優れていて、帯域幅を広くすると、音声の明瞭度と自然さが増すだけでなく、話す人を識別しやすくなる。このように広帯域音声コーデイングは次世代電話システムにとって重要な課題である。更に、マルチメディア分野が非常に成長したので、音楽や非音声信号を電話システムにより高品質で伝送することが望ましい。

高忠実度の線形PCM信号は、ビットレートが知覚エントロピーに関して非常に効率が悪い。CDの例では44.1kHzのサンプリング周波数と、サンプル当たり16ビットの分解能と、ステレオである。これは1411キロビット/秒のビットレートに等しい。ビットレートを大幅に下げたため、分割帯域知覚オーディオコーデックを用いて原始コーデイングを行うことができる。これらの自然オーディオコーデックは信号内の知覚無関係性と統計的冗長度を用いる。最高の

小しても実際上は劣化したと感ぜない。このように、ステレオでは約9.6キロビット/秒、すなわち約15:1の圧縮率で、非常に高い音質が得られる。或る知覚コーデックは更に高い圧縮率を用いる。このためには、サンプリングレート(したがってオーディオ帯域幅)を下げるのが普通である。また量子化レベルの数を減らし(量子化歪みが聞こえることがある)、また強化コーデイングによるステレオフィールドの劣化を用いるのが普通である。このような方法を余り用いると、耳障りな知覚劣化を生じる。現在のコーデック技術は飽和点に近く、符号化と利得が更に進むことは期待できない。符号化性能を高めるには、新しい方式が必要である。

人の声や鈴などの楽器は、振動システムから発生する調定常信号を生成する。フーリエ理論によると、周期的信号は周波数1, 2f, 3f, 4f, 5fなどの正弦波の和で表される。ただし、fは基本周波数である。これらの周波数は調定常系列を形成する。この信号の帯域幅を制限することは、調定常系列を切り捨てることに相当する。切捨てを行うと楽器や音声の音色が変わり、オーディオ信号は「弱い」または「鋭い」音になり、明瞭度が下がる。音質の主観的印象にとって高周波はこのように重要である。

従来の方法は、音声コーデック性能を高めることが主体で、特に音声符号化における問題である高周波再生(HFR)を目的としている。従来の方法は広帯域直線周波数ソフトや、非線形性や、エイリアシングを用いて[米国特許番号第5,127,054号]相互変調やその他の非調定常周波数成分を生成するので、これを音楽信号に適用するとひどい不協和音を生じた。この不協和音を音声符号化関係の文献では「耳障り」または「調子はずれ」の音と呼ぶ。他の合成音声HFR法は基本ピッチ推定に基づく正弦高調波を生成するので、定常音に限られる[米国特許番号第4,771,465号]。これらの従来の方法は低品質の音声応用には有用であるが、高品質音声または音楽信号には使えない。高品質のオーディオ原始コーデックの性能を高める方法がいくつかある。その1つは、デコーダで生成された合成雑音信号を用いて、以前はエンコーダで捨てられていた音声または音楽内の雑音的信号に代える(「雑音代替によるオーディオコーデックの改

コーデック技術を用いると、標準のCDフォーマット信号のデータを約90%縮

(Improving Audio Codes by Noise Substitution)」、D. Schulz, JAES, Vol. 44, No. 7/8, 1996)。これは雑音信号があるときに、本来正常に伝送される高帯域内で断続的に行われる。別の方法は、符号化の過程で失われた或る高帯域の高周波を再現する（「オーディオスペクトルコーダ(Audio Spectral Coder)」、A. J. S. Ferreira, AES Preprint 4201, 100th Convention, May 11-14, 1996 (Copenhagen)）。これも音信号とビット抽出に依存する。この2つの方法は低いデューティサイクルで動作し、比較的限定された符号化または性能の利得が得られる。

発明の概要

本発明はデジタル原始コーディング装置を実質的に改善する、より特定するとオーディオコーディングを改善する、新しい方法と装置を提供する。目的は、ビットレートの下下、または知覚品質の向上、またはその両方を含む。本発明は周波元量度を活用した新しい方法により、伝送または記憶を行う前に信号の通過帯域を縮減する可能性を提供する。本発明によりコーディングが高品質のスペクトル複製を行う場合は、知覚劣化は起こらない。複製ビットにおいて低帯域情報の符号化に更に多くのビットを割り当てて、より高い知覚品質を得ることができる。

本発明は、複製された周波系列は低帯域スペクトル成分と高帯域スペクトル成分の間の直接の関係に基づいて伸張することができると仮定する。この伸張された系列は、次の原則に従っていれれば初めの系列と似ていると知覚される。第1に、不協和音に属する人工音が出ないようにするために、外挿されるスペクトル成分は複製された周波系列と周波的に関係してないなければならない。本発明はスペクトル複製プロセスの手段として置換を用いる。これは確かにこの判定基準を満たす。しかし、置換した動作をするためには低帯域スペクトル成分が周波系列を形成する必要はない。その理由は、低帯域成分と周波的に関係する新しい複製成分は信号の雑音的または過渡的な性質を変えないからである。置換とは、部分音の周波数出を保持しながら部分音を音階上の1つの位置から別の位置に移すことである。第2に、複製された高帯域のスペクトル包絡線（すなわち、粗いスペクトル分布）は初めの信号と十分似ていなければならない。本発明は2つの動作モード

S

BR-1とSBR-2を提供する。この2つは、スペクトル包絡線を調整する方法が異なる。

SBR-1は中間品質コーディング応用を改善するシングルエンド配のプロセスであって、デコーダが受ける低帯域信号すなわち低域信号に含まれる情報に完全に依存する。この信号のスペクトル包絡線は、例えば多項式と規則の集合すなわちコードブックを用いて決定され、外挿される。この情報を用いて、複製された高帯域を絶えず調整し等化する。このSBR-1法は後処理の利点を持つ。すなわちエンコーダ側では修正する必要がない。放送業者はチャネルの利用度を高め、または知覚品質を高め、またはその両者が得られる。既存のビットストリーム構文と標準を修正せずに用いることができる。

SBR-2は高品質コーディング応用を改善するダブルエンド形のプロセスであって、SBR-1により伝送される低帯域信号の他に、高帯域のスペクトル包絡線を符号化して伝送する。スペクトル包絡線の変動速度は高帯域信号成分よりかなり低いので、限られた量の情報だけを伝送すればスペクトル包絡線を十分表すことができる。SBR-2を用いれば、既存の構文やプロトコルを全くまたは殆ど修正せずに現在のコーディング技術の性能を高めることができるので、今後のコーディングの開発の貴重なツールである。

SBR-1もSBR-2も、音響心理学モデルにより規定されたエンコーダがビット欠乏状態の下で低帯域の小さな通過帯域を停止したとき、これらを複製するのにも用いられる。低帯域内のスペクトル複製と低帯域外のスペクトル複製により、知覚品質が高まる。更に、SBR-1とSBR-2はビットレートスケールビリティを用いるコーディングにも用いることができる。この場合、受信器での信号の知覚品質は伝送チャネルの状態によって変わる。通常は、これは受信器でのオーディオ帯域幅の厄介な変動を意味する。この状態でSBR法を用いると常に高い帯域幅を保持するので、やはり知覚品質を高めることができる。

本発明は連続的に動作し、どんな種類の信号内容、すなわち音または非音（雑音的信号や通過信号）も複製する。更に、本発明のスペクトル複製法はデコーダ

で利用できる周波数帯域から、既知された帯域を知覚的に正確に複製することができ、したがってSBR法を用いると、従来の方法に比べて実質的に高いレベル

で符号化利得が得られ、または知覚品質を高めることができる。本発明を従来のコーデック改善法と組み合わせることができるが、組み合わせても性能が高まることは期待できない。

SBR法は次のステップを含む。

- ・ 初めの信号から得た信号を符号化し、信号の周波数帯域を既知する。既知は符号化の前か途中に行い、第1信号を形成する。

- ・ 第1信号の復号中またはその後で、第1信号の周波数帯域を既知して第2信号を形成する。

- ・ スペクトル包絡線を調整する。

- ・ 復号された信号と第2信号を組み合わせ出力信号を形成する。

第2信号の通過帯域は第1信号の通過帯域と重ならないようにまたは部分的に重なるように設定してよく、初めの信号および/または第1信号の時間特性、または伝送チャネルの状態に従って設定する。スペクトル包絡線の調整は、前記第1信号から初めのスペクトル包絡線を推定したもの、または初めの信号の伝送された包絡線情報に基づいて行う。

本発明は2つの基本型のトランスフォーマー（置換装置）を含む、すなわち、多帯域トランスフォーマーと時変パターン探索予測トランスフォーマーであって、これらは異なる特性を有する。本発明では基本的な多帯域置換を次のように行う。

- ・ 置換される信号を、それぞれ周波数 $[f_1, \dots, f_N]$ を含む通過帯域を持つ N (≥ 2) 個の通過帯域フィルタの集合で濾波して、 N 個の帯域信号を形成する。

- ・ 帯域信号の周波数を周波数 $M[f_1, \dots, f_N]$ を含む領域にシフトする。ただし、 $M \neq 1$ は置換係数である。

- ・ シフトされた帯域信号を結合して置換信号を形成する。

または、本発明ではこの基本的な多帯域置換を次のように行う。

- ・ 置換される信号を、低域型の実数値または複素値サブバンド信号を生成する

性質の分析フィルタバンクまたは置換を用いて帯域濾波する。

- ・ 任意のチャネル数 k の前記分析フィルタバンクまたは置換を、合成フィルタバンクまたは置換内の Mk ($M \neq 1$) チャネルに接続する。

- ・ 合成フィルタバンクまたは置換を用いて、置換された信号を形成する。

本発明の1つの改善された多帯域置換は位相調整を含み、基本的な多帯域置換の性能を強化する。

本発明では時変パターン探索予測置換を次のように行う。

- ・ 第1信号の通過検出を行う。
- ・ 通過検出の結果に従って、第1信号の一部を複写/複製すると共に、第1信号のどのセグメントを用いるかを決定する。

- ・ 通過検出の結果に従って、状態ベクトルとコードブック特性を調整する。
- ・ 前の同期点探索で見出された同期点に基づいて、第1信号の選択されたセグメント内の同期点を探す。

本発明のSBR法は次の特徴を有する。

- 1. この方法と装置はスペクトル領域内の信号元量性の新しい概念を活用する

- 2. この方法と装置は任意の信号に適用することができる。

- 3. 各周波集合は個々に作成して制動することができる。

- 4. 全ての複製された置換は既存の置換系列の延長を形成するようにして生成する。

- 5. スペクトル複製プロセスは置換に基づくもので、人工音は全くまたは殆ど知覚されない。

- 6. スペクトル複製は多数の小帯域および/または広い周波数範囲をカバーすることができる。

- 7. SBR-1法では、処理はデコード側だけで行う。すなわち、全ての標準およびプロトコルを修正せずに用いることができる。

- 8. SBR-2法は修正を全くまたは殆どせずに、殆どの標準やプロトコルに従って実現することができる。

9. SBR-2法はコーデック設計者に新しい強力な圧縮ツールを提供する。
10. 符号化利得は顕著である。
最も魅力的な応用は、各種の低ビットレートコーデック、例えば、MPEG1
／2層1／11／111 [米国特許番号5,040,217号]や、MPEG
2／4 AAC, Dolby AC-2／3, NTT TwinVQ [米国特許

- かの時間領域トランスフォーマーの実現のブロック図である。
図9a-図9cは、2次高調波を生成するよう構成された本発明のSTFT分
析および合成用の装置を表すブロック図である。
図10a-図10bは、本発明のSTFT装置内の直線周波数シフトを持つ1
つのサブバンドのブロック図である。
図11は、本発明の位相乗算器を用いる1つのサブバンドを示す。

番号第5,684,920号]や、AT&T/Lucent PACなど、の改
善に関する。。またこの発明は知覚品質を高めるための、高帯域CELPやSB
-ADPCM G.722などの、高品質音声コーデックにも有用である。上記
のコーデックはマルチメディアや、電話通話や、インターネットや、専門的な
用に広く用いられている。T-DAB (地上デジタルオーディオ放送) システ
ムは低ビットレートプロトコルを用いており、本方法を用いるとチャネル使用
度が上がり、またはFMやAM DABの品質を高めることができる。衛星S-
DABはシステムコストが非常に高いので、本方法を用いてDABマルチプレク
スのプログラムチャネル数を増やすことにより大きな利益を得る。更に、低ビ
ットレート電話モデムを用いて、インターネットにより初めて全帯域幅オーディ
オ実時間ストリーミングを達成することができる。

図面の簡単な説明

以下に本発明について添付の図面を参照して例を用いて説明するが、これは本
発明の範囲や精神を制限するものではない。

- 図1は、本発明の符号化装置内に挿入されたSBRである。
図2は、本発明の上部高調波のスペクトル複製を示す。
図3は、本発明の帯域内高調波のスペクトル複製を示す。
図4は、本発明のトランスフォーマーの時間領域乗算のブロック図である。
図5は、本発明のバターン探索予測トランスフォーマーの動作のサイクルを表す流
れ図である。
図6は、本発明の同期点の探索を表す流れ図である。
図7a-図7bは、本発明の過渡状態中のコードブック位置決めを示す。
図8は、本発明のSBR動作のための、適当なフィルタバンクに関するいくつか

好ましい実施の形態の説明

実施の形態の説明を通じて、自然オーディオ原始コーディング応用に重点を置いて述べる。しかし理解されるように、本発明はオーディオ信号の符号化や復号の応用の他に広範囲の原始コーディング応用に適用できるものである。

置換の基礎

本発明で述べる置換はスペクトル複製の理想的な方法であって、従来の方法に比べていくつもの大きな利点を持つ。すなわち、ビット抽出は必要なく、単一ビットで多音のプログラム材料において同じ高性能が得られ、置換は音信号にも非

音信号にも同じように良く適用できる。他の方法とは異なり、本発明の置換は任意の信号の種類において任意のオーディオ原始コーディング装置に用いることができる。

時変振幅を持つコサインの和の形の離散時間信号 $x(n)$ の正確な置換係数 M は、次の関係で定義される。

$$x(n) = \sum_{i=1}^N c_i(n) \cos(2\pi f_i n / f_s + \alpha_i) \rightarrow \quad (1)$$

$$y(n) = \sum_{i=1}^N c_i(n) \cos(2\pi f_i n / f_s + \beta_i) \quad (2)$$

ただし N は正弦波の数 (以後は部分音と呼ぶ)、 f_i と α_i はそれぞれ i 番目の入力周波数と時間包絡線と位相定数、 β_i は任意の出力位相定数、 f_s はサンプリング周波数、そして $0 \leq M f_i \leq f_s/2$ である。

図2に M 次高調波の生成を示す。ただし、 M は整数 ≥ 2 である。 M 次高調波という語は簡単のために用いた。実際は、このプロセスは或る周波数領域で全ての信号に M 次高調波を生成するが、多くの場合は次数の分らない高調波である。周波数領域の表現 $X(f)$ を持つ入力信号 201 の帯域は 0 から f_{max} の範囲に限定される。 f_{max}/M から $Q f_{max}/M$ (Q は望ましい帯域幅伸張係数で $1 < Q \leq M$) の範囲内の信号内容を帯域フィルタで取り出して、スペクトル $X_{BP}(f)$ を持つ帯域信号 203 を形成する。帯域信号を係数 M で置換して、 f_{max} から $Q f_{max}$ の範囲をカバーするスペクトル $X_T(f)$ を持つ帯域信号 205

を形成する。この信号のスペクトル包絡線をプログラム制御の帯域器で調整して、スペクトル $X_T(f)$ を持つ信号 207 を形成する。次にこの信号と遅延させた入力信号とを結合して、帯域フィルタとトランスポーズにより生じる遅れを補償して、 0 から $Q f_{max}$ の範囲をカバーするスペクトル $Y(f)$ を持つ出力信号 209 を形成する。または帯域幅は、逆時間変数は、逆時間変数 f_{max} と $Q f_{max}$ を用いて置換 M の後で行う。多重トランスポーズを用いて、異なる周波数成分を同時に生成することはもちろん可能である。図3に示すように入力信号が f_0 から $Q f_0$ にわたる通過帯域 301 を有する場合、上記の方式を用いて入力信号内の通過帯域を「充満する」こともできる。この場合は通過帯域 $[f_0/M, Q f_0/M]$ を取り出し

(303)、係数 M で $[f_0, Q f_0]$ に置換し (305)、包絡線を調整し (307)、遅延入力信号と結合してスペクトル $Y(f)$ を持つ出力信号 309 を形成する。

正確な置換の近似を用いてもよい。本発明では、このような近似の質を不協和音理論を用いて決定する。不協和音の判定基準は P_{IOMP} により示されており (f 音の協和と臨界帯域幅 (Tonal Consonance and Critical Bandwidth))、R. Plomp, W. J. M. Levelt JASA, Vol. 38, 1965]、2つの部分音の周波数の差がそれらが存在する臨界帯域の帯域幅の約 5% 乃至 50% 以内である場合は、2つの部分音は不協和と見なされる。例えば、所定の周波数の臨界 (critical) 帯域幅は次式で近似できる。

$$cb(f) = 25 + 75(f + 1) \left(\frac{f}{1000} \right)^{0.69} \quad (3)$$

ただし f と $c b$ は Hz で表す。更に P_{IOMP} は、2つの部分音の周波数の差がそれが存在する臨界帯域幅の約 5% 以下である場合は、人の聴覚システムはこの2つを識別することができないと述べている。式2の正確な置換を次式で近似することができ

$$y_{approx}(n) = \sum_{i=1}^N c_i(n) \cos(2\pi f_i n / f_s + \beta_i) \quad (4)$$

ただし、 Δf_i は正確な置換からの偏差である。入力部分音が周波数系列を形成する

場合は、本発明の仮説によると、置換された部分音の周波数系列からの置換はそれらが存在する限界帯域幅の5%を超えてはならない。従来の方法が不収な「耳障り」や「聞きはずれ」を生じるのは、広帯域連続周波数シフトにより許容できないほど大きな置換を生じるからである。従来の方法が1入力部分音に対して以上の部分音を生成するとき、1部分音として知覚されるためには部分音はやはり上述の置換限界内になければならぬ。非線形性などを用いる従来の方法は置換限界内にない相互変調部分音を作るので、良い結果が得られない。

本発明のスペクトル複製法に基づく上記の置換を用いると、次の重要な性質が得られる。

- ・ 通常は、複製された高周波と既存の部分音の間に周波数領域の重なりが起こらない。

らない。

- ・ 複製された部分音は入力信号の部分音と周波数的に関係があり、耳障りな不協和音すなわち人工音を一切生じない。
- ・ 複製された高周波のスペクトル包絡線は入力信号スペクトル包絡線の滑らかな相似を形成し、初めの包絡線と知覚的に一致する。
- ・ 時変パターンの探索予測に基づく置換

必要なトランスポートを設計するには種々の方法がある。一般的な時間領域表現では、ピッチ周期に基づいて信号セグメントを複製することにより信号を時間的に伸張する。次にこの信号を異なる速度で読み出す。残念ながらこの方法は、信号セグメントを正確に時間接続するのにピッチ検出に厳密に依存する。更に、ピッチ周期に基づく信号セグメントには制約があるので、置換信号に特徴になる。検出されたピッチ周期は実際の置換信号よりかなり長いことがあるので、全周波信号を単に時間的に伸張するのではなく複製するという危険が大いことは明らかである。別の種類の時間領域/圧縮を得る[「音声のパターン探索予測(Pattern Search Prediction of Speech)」, R. Bognier, T. Li, Proc. ICASSP, '89, Vol. 1, May 1989や、「非線形共振器モデルに基づく音声のタイムスケール修正(Time-Scale Modification of Speech based on a nonlinear Oscillator Model)

J. G. Kubin, W. B. Kleijin, IEEE, 1994]。これは粒状合成(granular synthesis)の1つの形であって、入力信号を小さな部分(細粒)に分割し、これを用いて出力信号を合成する。この合成には通常は信号セグメントの相関を取って、最良の接点を選択する。これは、出力信号を形成するのに用いるセグメントがピッチ周期に依存しないのでピッチ検出という厄介なタスクを必要としないことを意味する。しかしこの方法にも信号振幅が急速に変わるときの問題がまだあり、高品質の置換のためには高速計算が必要になる。しかし改善された時間領域のピッチシフト/トランスポートを用いることにより、定常音(音または非音)でも過渡音でもシステムパラメータを用いることにより、定常音(音または非音)でも過渡音でも、高い置換品質の一層正確な置換を低い計算コストで行うことができる。

次に図面を参照する。各図面の同じ要素は同じ番号で示す。図4に9個の別個

のモジュールを示す。すなわち、過渡検出器401と、窓位置調整器403と、コードブック発生器405と、同期信号選択器407と、同期位置メモリ409と、最小遅延器411と、出力セグメントメモリ413と、混合ユニット415と、ダウンサンプリング417である。入力信号はコードブック発生器405と過渡検出器401に入る。過渡信号を検出すると、その位置を窓位置モジュール403に送る。このモジュールは窓の大きさや位置を規定し、コードブックを作る。入力信号と掛け算する。別のトランスポートに接続している場合は、コードブック発生器405は同期選択モジュール407から同期位置を受ける。この同期位置はコードブック内にある場合は、これを用いて出力セグメントを生成する。ない場合は、コードブックを最小遅延器411に送り、最小遅延器411は新しい同期位置を送る。新しい出力セグメントと前の出力セグメントを共に混合モジュール415で窓に入れ、モジュール417でダウンサンプリングする。説明を明確にするために非線形空間表現を用いる。ここで状態ベクトルすなわち詳細は、入力信号と出力信号を表す。入力信号を次の状態ベクトル $x(n)$ で表す。

$$x(n)=[x(n), x(n-D), x(n-2D), \dots, x(n-(N-n)D)]$$

これは入力信号のN個の遅延サンプルから得られる。ただし、Nは状態ベクトルの次元、Dはベクトルを作るのに用いる入力サンプル間の遅れである。状態マッピングにより各状態ベクトル $x(n-1)$ の後のサンプル $x(n)$ が得られる。これを式6で表す。ただし、 $a(*)$ はマッピングである。

(6)

$$x(n) = a(x(n-1)).$$

本方法では、状態遷移コードブックを用いて、状態マッピングにより前の出力に基づいて次の出力を決定する。草しのコードブックは絶えず再構築され、状態ベクトルと各状態ベクトルに続く次のサンプルを含む。各状態ベクトルはその隣接状態ベクトルからKサンプル離れている。これによりこの装置は、現在処理中の信号の特性に従って時間分解能を調整することができる。ただし、Kは最大分解能を表すものに等しい。コードブックを作成するのに用いる入力信号セグメントは、起こり得る過渡信号の位置と前のコードブック内の同期位置に基づいて選択する。

これは理論的には、マッピング $a(*)$ はコードブックに含まれる全ての遷移について評価することを意味する。

(7)

$$a\left(\begin{bmatrix} x(n-L) \\ \vdots \\ x(n-L+K) \\ \vdots \\ x(n-D) \end{bmatrix}\right) = \begin{bmatrix} x(n-L+1) \\ \vdots \\ x(n-L+K+1) \\ \vdots \\ x(n) \end{bmatrix}.$$

この遷移コードブックを用いて、現在の状態ベクトル $y(n-1)$ に最も近いコードブック内の状態ベクトルを探して新しい出力 $y(n)$ を計算する。この最も近い隣接状態ベクトルの探索は最小値を計算することにより行い、新しい出力サンプルを得る。すなわち、

(8)

$$y(n) = a(y(n-1)).$$

しかしこの装置はサンプル毎に作用するよう制限されているわけではなく、好ましくはセグメント毎に作用する。新しい出力セグメントを窓に入れ、前の出力セ

グメントと加算し混合した後、ダウンサンプリングする。ピッチ置換係数は、コードブックで表される入力セグメントの長さ、コードブックから読み出される出力セグメントの長さの比で決まる。

図面に戻って、図5と図6はトランスポーズの動作のサイクルを示す流れ図である。501に入力データが入り、503で入力信号のセグメントの過渡検出を行う。過渡信号の探索は出力セグメントの長さに等しいセグメントの長さについて行う。505で過渡信号が見つかったと、507で過渡信号の位置を配信し、509でパラメータL(コードブックの長さを表す)と、K(サンプル内の各状態ベクトル間の距離を表す)と、D(各状態ベクトル内のサンプル間の遅れを表す)を調整する。511で過渡信号の位置と前の出力セグメントの位置を比較して、過渡信号の処理が済んだかどうか判定する。513で処理が済んだ場合は、515でコードブックの位置(窓L)と、パラメータKとLとDを調整する。必要なパラメータ調整が終わると、517で過渡検出の結果に基づいて新しい同期

点をすなわち接続点を探す。この手続きを図6に示す。601でまず前の同期点に基づいて、次式により新しい同期点を計算する。

(9)

$$Sync_pos = Sync_pos_old + SM - S,$$

ただし、 $Sync_pos$ と $Sync_pos_old$ はそれぞれ新および旧の同期位置、Sは処理した入力セグメントの長さ、Mは置換係数である。603でこの同期点を用いて新しい接続点の精度と旧い接続点の精度を比較する。605でこれが前と同じまたは一層良く一致していて、しかもコードブック内にある場合は、607で新しい同期点を返す。一致が良くない場合は、609で新しい同期点をループで探す。これは類似性測度(この場合は611の最小遊離度)を用いて行うが、時間領域または周波数領域の相関を用いることも可能である。613で前の位置より良く一致する位置がある場合は、615で同期位置を配信する。617で全ての位置を調べると、619でこのシステムは図5の遅れ図に戻る。得られた新しい同期点を519で配信し、521でコードブックから新しいセグメントを、所定の同期点から順番に読み出す。523でこのセグメント

を窓に入れて前のセグメントと加算し、525で整数係数によりダウンスamplingし、527で出力バッファに記憶する。

図7は、コードブックの位置に関して過渡状態にあるシステムの行動を示す。過渡信号の前には、入力セグメント1を表すコードブック1は、セグメント1の「左に」位置する。相関セグメント1は前の出力の一瞬を差し、これを用いてコードブック1内の同期点1を見つける。過渡信号を送出して過渡信号の点进行处理すると、コードブックを図7aに従って動き、現在処理中の入力セグメントのびコードブックの「右に」なるまで静止する。このシステムは過渡信号より前の同期点を探索することができないので、これにより過渡信号を復号することは不可能になる。

パターン探索予測に基づく多くのビットトランスポーズまたは時間エキスパンダは音声および単一ビット材料に満足できる結果を与える。しかし音楽のような高度に複雑な信号では、特に整数係数が大きい場合は、その性能は急速に悪化する。本発明は性能を高めるいくつかの解決法を提案するので、どんな種類の信号

。これについては後の実施例で示す。

すでに述べた時間領域トランスポーズを用いて、以下の例のようにSBR-1とSBR-2を実現する。これは例示であって制限するものではない。図8では、3つの時間伸張モジュールを用いて2次、3次、4次の高周波を生成する。この例では各時間領域伸張/トランスポーズは広帯域信号に作用するので、整数の後では手段がないことを考慮して、別の帯域分割器を追加せずに整数の前に原始周波数範囲のスペクトル包絡線を調整するといふ。スペクトル包絡線調整器801、803、805はそれぞれいくつかのフィルタバンクチャンネルに作用する。包絡線調整器内の各チャンネルの利得は、整数後に出力での和813、815、817が所望のスペクトル包絡線を生成するように設定しなければならない。トランスポーズ807、809、811は相互に接続され、同期位置情報を共有する。これは、或る条件下では、別個の整数ユニット内で相関中にコードブック内で見出される同期位置の間に高い相関が起こる、という事実に基づいている。やはり例であって本発明の範囲を制限するものではないが、4次の周波トランスポーズは2次の周波トランスポーズに比べて、時間フレームでは半分で動作するがデュー

ーティサイクルでは2倍で動作すると仮定する。更に、2つのエキスパンダに用いられるコードブックは同じと仮定し、2つの時間領域エキスパンダの同期位置をそれぞれsync_pos4とsync_pos2で表すと次の関係がある。

$$sync_pos2 = sync_pos4 - n \cdot 4 \cdot S - sync_offset, \quad (n=1,2,3,4, \dots) \quad (10)$$

ただし、

$$sync_offset = sync_pos4 - sync_pos2, \quad (n=0) \quad (11)$$

またSはコードブックで表される入セグメントの長さである。どちらの同期位置ポインタもコードブックの終わりに到達していなければこれは有効である。通常の動作中は、2次周波トランスポーズが処理する時間フレーム毎にnは1ずつ増加し、ポインタのどちらかが結局コードブックの終わりに到達すると、カウンタn=n=0に設定し、sync_pos2とsync_pos4を個別に計算する。4次の周波トランスポーズに接続すると、3次の周波トランスポーズにつ

いても同じ結果が得られる。
上に述べたようにいくつかの相互接続された時間領域トランスポートを高次の高周波の生成に用いると、計算が大層に減る。更に、ここに述べたように時間領域トランスポートを適当なフィルタバンクと共に用いると、生成されたスペクトルの包絡線を調整することができ、しかも時間領域トランスポートの簡単な低い計算コストを保つことができる。それは、これらが多少でも固定点計算と加算/減算の演算だけを用いて実現できるからである。
例示であって制限的でないこの発明の他の例は、
・ サブバンドフィルタバンク内の各サブバンド内で時間領域トランスポートを用いて、各トランスポートの信号の複雑さを減らす。
・ 時間領域トランスポートと周波数領域トランスポートを共に用いて、処理中の入力信号の特性に従ってシステムが異なる電換法を用いることができるようにする。
・ 広帯域音声コーディング内で時間領域トランスポートを用いて、例えば直線予測の後に得られる残留信号に作用する。

$$X_k(n) = \sum_{p=0}^{N-1} x(p) h(n-p) e^{-j\omega_k p} \quad (12)$$

ただし、 $k=0, 1, \dots, N-1$ 、 $\omega_k = 2\pi k/N$ 、 $h(n)$ は窓である。
窓が次の条件

$$\begin{cases} h(0)=1 \\ h(n)=0 \end{cases} \quad (n=2N \pm 2M \pm 3N, \dots) \quad (13)$$

を満たす場合は逆変換が存在して次式で与えられる。

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X_k(n) e^{j\omega_k n} \quad (14)$$

正変換は分析器と見なしてよく (図9a参照)、インパルス応答 $h(n)$ $e^{j\omega_k n}$ $p(j\omega_k n)$ 901 を持つ N 個のBPFフィルタのバンクの後に、キャリア $e^{j\omega_k n}$ $p(-j\omega_k n)$ 903 を持つ N 個の乗算器のバンクがあり、BP信号を $0Hz$ 付近の領域にシフトダウンして、 N 個の分析信号 $X_k(n)$ を形成する。窓はプロタイプLPFフィルタとして動作する。 $X_k(n)$ は小さな帯域幅を持ち、通常は 905 でダウンサンプリングされる。したがって式 12 は $n=rR$ のときだけ計算すればよい。ただし R は 10 進化係数、 r は新しい時間変数である。 $X_k(n)$ はアップサンプリングにより $X_k(rR)$ から回復することができる (図

9b参照)。すなわち、 907 でゼロを挿入した後、 909 でLP滤波する。逆変換は合成器と見なされ、その構成は、信号 $X_k(n)$ を初めの周波数にまでシフトするキャリア $(1/N) e^{j\omega_k n}$ 911 を持つ N 個の乗算器のバンクの後に、全てのチャネルからの貢獻 $y_k(n)$ を加算する段 913 (図9c参照) がある。STFTおよびISTFTを再配列してDFTおよびIDFTを用いてよく、これによりFFTアルゴリズムを用いることができる [「高速フーリエ変換を用いた位相コーダの実現 (Implementation of the Phase Vocoder using the Fast Fourier Transform)」, M. R. Portnoff, IEEE ASSP, Vol. 24, No. 3, 1976]。

図9cは $N=3$ を持つ2次高調波 ($M=2$) を生成するパッチ 915 を示す。簡単のために、チャネル 0 乃至 16 だけを示す。BP 16 の中心周波数はす

認識すべきことは、上に述べた方法は、サンプリレート変換を単に省略してタイムスケール修正に用いるときだけ使われているということである。更に理解すべきことは、上述の方法はより高いピッチへのピッチ電換 (すなわち時間伸張) に焦点を当てているが、当業者に明らかにならないように、同じ原理はより低いピッチへの電換 (すなわち時間圧縮) にも適用できることである。

フィルタバンクを用いた電換

種々の新しい革新的なフィルタバンクを用いた電換技術について以下に説明する。電換される信号を一通のBP信号またはサブバンド信号に分割する。次にサブバンド信号を正確にまたは近似的に電換する。これを行うには、分析サブバンドと合成サブバンドを再接続する (以後、「パッチ」と呼ぶ) といふ。この方法について、まず短時間フーリエ変換 (STFT) を用いて説明する。

離散時間信号 $x(n)$ の N 点STFTを次のように定義する。

イキスト周波数に等しく、チャンネル17乃至31は負の周波数に相当する。P917および利得ブロック919で示すブロックについては後で説明するので、とりあえずは除いて考える。この例の入力信号は帯域が制限されており、チャンネル0乃至7だけが信号を含む。分析チャンネル8乃至16は空であって、合成器にマッピングする必要はない。分析チャンネル0乃至7は、入力信号選択路に相当する合成チャンネル0乃至7に接続する。また分析チャンネルk (4 ≤ k ≤ 7) は合成チャンネルMk (M=2) にも接続し、BPフィルタkの中心周波数の2倍の周波数領域に信号をシフトする。したがって、信号は初期の範囲にアップリングされるだけでなく、1オクターブ上に置換される。周波生成を逐次値フィルタ応答および変調器に關して開くには、負の周波数も考慮しなければならぬ(図10aの下側の分岐を参照)。したがって、再マッピングk→Mk+1001とN-k-N-Mk+1003 (4 ≤ k ≤ 7) の組み合わせ出力を評価しなければならぬ。

これから次が得られる。

$$x(n) = \frac{2}{N} \left[r(n) \cdot h(n) \cos(\omega_k n) \right] \cos(M - 1) \omega_k n + \frac{2}{N} \left[r(n) \cdot h(n) \sin(\omega_k n) \right] \sin(M - 1) \omega_k n \tag{15}$$

ただし、M=2である。式15は、入力信号のBP諸波の後に、直交周波数シフ

トすなわち上側波帯(USB)波調、すなわち上側波帯を用いる直交波帯波調(図10b参照)が視くと考えてよい。ただし1005と1007はヘルベルト変成器を形成し、1009と1011はコサインおよびサインキャリアを持つ乗算器であり、1013は上側波帯を選択する選分器である。明らかに、このような多帯域BPおよびSSB法は明示的に、すなわちフィルタバンクパッチングなしに、時間領域または周波数領域で実現され、個々の通過帯域と変換器周波数を任意に選択することができる。

式15では、分析チャンネルkの通過帯域内の周波数ω_kを持つ正弦波は周波数Mω_k+ (ω_i-ω_k)で調波を生成する。したがって、基本的多帯域変換と呼ぶ方法だけが、周波数ω_i=ω_k (4 ≤ k ≤ 7)を持つ入力信号の正確な高調波を

生成する。しかし、フィルタの数が十分大きい場合は、正確な置換からの偏差は無視できる(式4参照)。更に、各分析チャンネルが最大1部分音を含む場合は、P917(図9c)で示すブロックを挿入することにより任意の周波数の選定、部分音時は正確に置換される。この場合、X_k(rR)は、部分音周波数ω_iと分析フィルタの中心周波数ω_kの差に等しい周波数を持つ複素指数関数である。正確な置換Mを得るには、これらの周波数は上の周波数関係を係数Mだけ増加させて、ω_i-Mω_k+M (ω_i-ω_k)=Mω_iに修正しなければならない。X_k(rR)の周波数はそれぞれの重ならない位相角の時間微分に等しく、連続した位相角の1次部分を用いて推定してよい。周波数の推定にMを掛けて、これらの新しい周波数を用いて合成位相角を計算する。しかし、位相定数を除いて、分析引数にMを直接掛けるという簡易法で同じ結果が得られ、周波数を推定する必要がなくなる。これを、ブロック917を表す図11に示す。まず、X_k(rR) (この例では4 ≤ k ≤ 7)を直交座標から極座標に変換する。これをブロックR→P、1101で示す。1103で引数にM=2を掛け、絶対値は変えない、次に1105で信号を変換して直交座標に戻して(P→R)信号YMK(rR)を形成し、図9cの合成チャンネルに与える。この改善された多帯域置換法には2つの段がある。基本的な方法と同様にパッチにより相い置換を行い、位相乗算器により微細な周波数訂正を行う。上の多帯域置換法はSTFTを用いる従来のピッチシフト法とは異なる。STFTでは合成にルックアップテーブル発振器を用いる。また

ISTFTを合成に用いる時は信号の時間を伸ばして10進化する、すなわちパッチを用いない。

図9cの調波パッチは、置換係数が2でない場合は容易に修正される。図12は3次高調波を生成するパッチ1203を示す。1201は分析チャンネル、1205は合成チャンネルである。異なる次数の高調波も図13に示すように同時に作ることもできる(ここでは2次と3次の高調波を用いる)。図14は、2次と3次と4次の高調波が重ならない組み合わせを示す。最低の可能な調波数をできるだけ高い周波数で用いる。調波Mの所先範囲の上限を超えると、調波M+1を

用いる。図15は、全ての合成チャンネル ($N=64$ 、チャンネル0-32を示す) にマッピングする方法を示す。非帯域指数を持つ全ての高帯域チャンネルは、原始チャンネル番号と祖先チャンネル番号の関数 $k_{dest} = M \cdot k_{source}$ に従ってマッピングする。ただし、 M は k_{source} が低帯域にあり k_{dest} が高帯域にあるという条件を満たす最小整数 ≥ 2 である。したがって、どの合成チャンネルも2つ以上の分析チャンネルから信号を受けない。多数の高帯域チャンネルは $k_{source} = 1$ または低帯域チャンネル $k_{source} > 1$ にマッピングしてよく、これは上述の関数のよい近似を作る (図15には $M=2, 3, 4, 5$ における非帯域接続だけを示す)。

異なる分析チャンネルからの振幅情報と位相情報を組み合わせることも可能である。振幅信号 $|X_k(r, R)|$ は図16のように接続してよい。位相信号 $a(r, g(X_k(r, R)))$ は図16の原理に従って接続する。このようにして、低帯域周波数を置換し、図2の置換から得られる伸張包絡線ではなく原始領域包絡線の周相的帰還しを生成する。「空」原始チャンネルを挿入しないようにするため、ゲートや他の手段を用いてよい。図17は別の応用であって、高位のサブバンドから低位のサブバンドへの接続を用いて、高帯域信号または低帯域信号に分離高周波を生成する。上記の置換を用いるとき、信号の特性に基づいてパッチの適応的替えを用いるとよい。

上の説明では、入力信号に含まれる最高周波数はナイキスト周波数よりかなり低いと仮定した。したがってサンプルレートを増やさずに帯域幅を拡張することできた。しかしこれはいつでもできるわけではなく、事前のアップサンプリングが必要である。置換にフィルタバンク法を用いるときは、アップサンプリングをプロセスに統合することが可能である。

多くの知覚コーデックは、時間から周波数へのマッピングに最大10進化フィルタバンクを用いる [「知覚コーディング入門(Introduction to Perceptual Coding)」, K. Brandenburg, AES, デジタルオーディオのビットレート減少に関する論文集(Collected Papers on Digital Audio Bitrate Reduction), 1996]。図18aは知覚エンコーディング装置の基本構造を示す。分析フィルタバンク1801

は入力信号をいくつかのサブバンド信号に分割する。1803で、減らしたビット数を用いてサブバンド信号を個別に量子化する。量子化レベル数は、最小マスキングしきい値を推定する知覚モデル1807から決定する。サブバンドサンプルを基礎化し、任意の冗長度符号化法で符号化し、1805で基礎化係数やビット割当て情報やその他のコーデック特定データから成る側情報と組み合わせ、直列ビットストリームを形成する。次にこのビットストリームを配位しまたは伝送する。図18bのデコーダでは、1809で符号化ビットストリームのデマルチプレクシングを行い、復号して、1811でサブバンドサンプルを等しい数のビットに再量子化する。1813で、合成フィルタバンクはサブバンドサンプルを組み合わせて初めの信号を再生する。最大10進化フィルタバンクを用いて実現すると計算コストが大幅に減る。以下の説明では、コサイン変調フィルタバンクに焦点を当てる。しかし理解すべきことは、本発明は、ウェーブレット変換のフィルタバンク解釈や、他の不等帯域幅フィルタバンクまたは変換や、多次元フィルタバンクまたは変換などの他の種類のフィルタバンクまたは変換を用いて実現することができることである。

例であって制限するものではないが、以下の説明ではLチャンネルコサイン変調フィルタバンクは入力信号 $x(n)$ をL個のサブバンド信号に分割すると仮定する。最大10進化フィルタバンクの一般構造を図19に示す。分析フィルタ $H_k(z)$ (1901 ($k=0, 1, \dots, L-1$)) で示す。1903でサブバンド信号 $y_k(n)$ を最大10進化する。各サンプル周波数は f_s/L である。ただし、 f_d は $x(n)$ のサンプリング周波数である。合成部では1905で内挿

し1907で減波した後、サブバンド信号を再組立てして $\hat{x}(n)$ を生成する。合成フィルタを $F_u(z)$ で示す。更に、本発明は $\hat{x}(n)$ にスペクトル複製を

行い、強化信号 $y(n)$ を生じる。

サブバンド信号をQチャンネルフィルタバンクで合成するときは、L個の低帯域チャンネルだけを用い、また帯域幅伸張係数QはQLが整数値になるように選択するが、この合成によりサンプリング周波数 Qf_s を持つ出力ビットストリームが得られる。したがって、拡大フィルタバンクはLチャンネルフィルタバン

クの後にアップサンブラがあるかのように動作する。この場合は $L(Q-1)$ 個の高帯域フィルタは用いない(ゼロを与える)ので、オーディオ帯域幅は変わらない。

フィルタバンクは単に $\hat{x}(n)$ をアップサンプリングしたものを再構築するだけである。しかし L 個のサブバンド信号を高帯域フィルタにパッチングした場合は、 $\hat{x}(n)$ の帯域幅は係数 Q だけ増えて $y(n)$ を生成する。これは本発明の基本的多帯域トランスフォーマの最大10進化フィルタバンク版である。この方式を用いると、アップサンプリングプロセスは前に説明した合成濾波に結合される。注意したいのは、どんな大きさのフィルタバンクを用いても、出力信号のサンプルレートの異なり、したがって帯域幅伸張係数は異なることである。整数

場合は、 $\hat{x}(n)$ の帯域幅は係数 Q だけ増えて $y(n)$ を生成する。これは本発

明の基本的多帯域トランスフォーマの最大10進化フィルタバンク版である。この方式を用いると、アップサンプリングプロセスは前に説明した合成濾波に結合される。注意したいのは、どんな大きさのフィルタバンクを用いても、出力信号のサンプルレートの異なり、したがって帯域幅伸張係数は異なることである。整数

置換係数 M を持つ本発明の基本的多帯域置換法を用いて $\hat{x}(n)$ にスペクトル複

素を行うには、次式でサブバンド信号をパッチングする。

$$y_{mk}(n) = e_{mk}(n)(-1)^{Qk-1}x_k(n), \quad (16)$$

ただし、 $k \in [0, L-1]$ であって $Mk \in [L, QL-1]$ になるように選択され、 $e_{mk}(n)$ は包絡線訂正、 $(-1)^{(Q-1)k}$ はスペクトル反転サブバンドの訂正係数である。スペクトル反転はサブバンド信号の10進化の結果であり、反転信号はこれらのチャンネル内の1つ置きのサンプルの符号を変えることにより再反転する。図20は16チャンネルの合成フィルタバンクであって、2009で置換係数 $M=2$ 、 $Q=2$ についてパッチングされている。ブロック2001と2003はそれぞれ図19の分析フィルタ $H_k(z)$ とデジメータである。同様に、2005と2007は補間回路と合成フィルタ $F_k(z)$ である。これにより、式16は、受信データの4つの上位周波数サブバンド信号を、合成フィルタバンク内の8つの最上位チャンネルの1つ置きのチャンネルにパッチングすること

に簡便化される。スペクトル反転を行ったので、1つ置きのパッチングされたサブバンド信号は合成する前に周波数を反転しなければならない。更に2011で、パッチングされた信号の振幅を $SBR-1$ または $SBR-2$ の原理に従って調整しなければならない。

本発明の基本的多帯域置換法を用いると、生成される高周波は一般に基本波の正確な倍数にならない。各サブバンドの最低周波を除く全ての周波数は正確な

置換とは或る程度異なる。更に、ターゲット間隔は原始間隔より広い周波数範囲をカバーするので、複製スペクトルはゼロを含む。更に、サブバンド信号の周波数はターゲット間隔に分離されるので、コサイン変調フィルタバンクのエリアス消し特性はなくなる。すなわち、隣接サブバンド信号は高帯域領域で重ならない。しかし、当業者に知られているエリアス削減法を用いればこの種の人工音を減らすことができる。この置換法の利点は、実現が容易なことと、計算コストが非常に低いことである。

正歪波を完全に置換するため、改善された多帯域置換法の効果的な最大10進化フィルタバンクを用いた解決法を以下に提示する。このシステムは追加の修正分析フィルタバンクを用い、合成フィルタバンクはVaidyanathanにより述べられている方法でコサイン変調する(「マルチレートシステムとフィルタバンク(Nutall, P. Vaidyanathan, Prentice Hall, Englewood Cliffs, New Jersey, 1993, ISBN 0-13-605718-7)」。最大10進化フィルタバンクに基づいて、本発明の改善された多帯域置換法を用いた操作のステップを図21の時図と、図22の流れ図で以下に示す。

1. L 個の受信サブバンド信号を QL チャンネルのフィルタバンク2101、2201、2203で合成して信号 $x_1(n)$ を形成する($L(Q-1)$ 上制チヤンネルにはゼロを与える)。したがって、信号 $x_1(n)$ は帯域幅伸張係数 Q でオーバーサンプリングされる。
2. 2103、2205で $x_1(n)$ を係数 Q でダウンサンプリングして信号 $x_2(n')$ を形成する。すなわち、 $x_2(n') = x_1(Qn')$ である。
3. 2207、2209、2211で、 $T = KM/Q$ で整数になるように整数値 K を合成フィルタバンクの大きさとして選択する。ただし、 T は修正された分

析フィルタバンクの大きさ、 M は置換係数である。好ましくは、 K は定常(寄)信号では大きく、動的(過渡的)信号では小さくなるように選ぶ。

4. 2107、2213で、 T チャンネルの修正された分析フィルタバンクで $x_2(n')$ を濾波し(T 分析フィルタは指数関数的に変調される)、複素値のサブバンド信号の集合を生成する。サブバンド信号を係数 T/M でダウンサン

リングし、サブバンド信号 $v_k(n)$ ($k=0, 1, \dots, T-1$) を生成する。したがって、フィルタバンクは係数 M でオーバーサンプリングされる。

5. サンプル $v_k(n)$ (n) を複素数表現 (振幅と位相角) に変換する。2109, 2215 で、位相角に係数 M を掛けて、サンプルを図 11 の方法で変換して直交座標表現に戻す。複素値サンプルの真数部を取り、信号 $s_k(n)$ を生成する。この操作の後、信号 $s_k(n)$ (n) を厳密にサンプリングする。

6. 2111, 2217 で、SBR-1 または SBR-2 の原理に従って信号 $s_k(n)$ (n) の利得を調整する。

7. 2105, 2221 で、サブバンド信号 $s_k(n)$ (n) (ただし、 $k \in [T/M, \min(K, T-1)]$) を通常のコサイン変調 K チャンネルフィルタバンクで合成して、チャンネル 0 乃至 $T/M-1$ にゼロを与える。これにより、信号 $x_3(n)$ を生成する。

8. 2223 で、最終的に $x_3(n)$ と $x_1(n)$ を加算して $y(n)$ を得る。これが所望のスペクトル複製信号である。

複製係数 M の異なる値についてステップ 3 乃至 6 を繰り返して、 $x_1(n)$ に多重化調整を加える。この動作モードを図 21 の点線で示し、また図 22 の 11-2219 のループの繰り返しで示す。この場合、 M の全ての選択値 (M は整数値) に対して T が整数値になるように K を選ぶ。好ましくは、 K/Q が正の整数になるように K を選ぶ。全てのサブバンド信号 $s_k(m)$ (n) (ただし、 $i=1, 2, \dots, m$ 、また m は複製係数の数) を、式

ただし、 $k=K/Q, K/Q+1, \dots, \min(K, T)-1$ ステップ 7 に従って、サブバンド信号 $s_k(n)$ を K チャンネルフィルタバンクで 1 度に変換する。

ステップ 4 の修正された分析フィルタバンクは、コサイン変調フィルタバンクの理論から得られる。ここで、変調重は合わせ変換 (MLT) [「効率的変換/サブバンド符号化のための重ね合わせ変換 (Lapped Transform for Efficient Transform/Subband Coding)」, H. S. Malvar, IEEE Trans ASSP, vol. 38, no. 6, 1990] は特殊なケースである。T チャンネルのコサイン変調フィルタバンク内のフィルタのインパルス応答 $h_k(n)$ は次のように書かれる。

$$h_k(n) = C p_k(n) \cos \left[\frac{\pi}{2T} (2k+1)(n - \frac{N-1}{2}) + \phi_k \right] \quad (19)$$

ただし、 $k=0, 1, \dots, T-1$, N は低域プロトタイプフィルタ $p_0(n)$ の長さ、 C は定数、 ϕ_k は隣接チャンネル間のエリアスを打ち消す位相角である。 ϕ_k の簡約は次式で表され、

$$\phi_k = \pm \frac{\pi}{4}, \phi_{k+1} = \pm \frac{\pi}{4} \text{ and } \phi_k = \phi_{k+1} \pm \frac{\pi}{2} \quad (20a-c)$$

これを簡単化すると次の閉じた形式表現になる。

$$\phi_k = k(-1)^k \frac{\pi}{4} \quad (21)$$

ϕ_k をこのように選ぶと、インパルス応答を持つ合成フィルタバンク

$$f_k(n) = C p_k(n) \cos \left[\frac{\pi}{2T} (2k+1)(n - \frac{N-1}{2}) - \phi_k \right] \quad (22)$$

を用いて、完全な再構成システムまたは近似的な再構成システム (疑似 QMF システム) が得られる。
次のフィルタを考える。

$$h'_k(n) = C p_k(n) \sin \left[\frac{\pi}{2T} (2k+1)(n - \frac{N-1}{2}) + \phi_k \right] \quad (23)$$

ただし、 $h'_k(n)$ はプロトタイプフィルタ $p_0(n)$ をサイン変調したもので

リングし、サブバンド信号 $v_k(n)$ ($k=0, 1, \dots, T-1$) を生成する。したがって、フィルタバンクは係数 M でオーバーサンプリングされる。

5. サンプル $v_k(n)$ (n) を複素数表現 (振幅と位相角) に変換する。2109, 2215 で、位相角に係数 M を掛けて、サンプルを図 11 の方法で変換して直交座標表現に戻す。複素値サンプルの真数部を取り、信号 $s_k(n)$ を生成する。この操作の後、信号 $s_k(n)$ (n) を厳密にサンプリングする。

6. 2111, 2217 で、SBR-1 または SBR-2 の原理に従って信号 $s_k(n)$ (n) の利得を調整する。

7. 2105, 2221 で、サブバンド信号 $s_k(n)$ (n) (ただし、 $k \in [T/M, \min(K, T-1)]$) を通常のコサイン変調 K チャンネルフィルタバンクで合成して、チャンネル 0 乃至 $T/M-1$ にゼロを与える。これにより、信号 $x_3(n)$ を生成する。

8. 2223 で、最終的に $x_3(n)$ と $x_1(n)$ を加算して $y(n)$ を得る。これが所望のスペクトル複製信号である。

複製係数 M の異なる値についてステップ 3 乃至 6 を繰り返して、 $x_1(n)$ に多重化調整を加える。この動作モードを図 21 の点線で示し、また図 22 の 11-2219 のループの繰り返しで示す。この場合、 M の全ての選択値 (M は整数値) に対して T が整数値になるように K を選ぶ。好ましくは、 K/Q が正の整数になるように K を選ぶ。全てのサブバンド信号 $s_k(m)$ (n) (ただし、 $i=1, 2, \dots, m$ 、また m は複製係数の数) を、式

$$s_i(n') = \sum_{k=0}^{M-1} s_k^{(M)}(n') \quad (17)$$

を用いて全ての適用可能な k について加算する。図 22 のループの第 1 輪通しでは、信号 $s_k(n')$ (ただし、 $k=0, 1, \dots, K-1$) はゼロだけのサ

ブバンドサンプルと考えてよい。全てのループにおいて、2219 で次式により新しいサンプルを $s_k(n')$ に加える。

$$s_i(n') = s_i(n') + s_i^{(M)}(n') \quad (18)$$

ある。フィルタ $H^*_k(z)$ と $H_k(z)$ は、同じ通過帯域支援を有するが位相応答が異なる。フィルタの通過帯域は英訳は相互のヒルベルト変換である（これは $\omega=0$ および $\omega=\pi$ に近い周波数では有効でない）。式19と式23を結合すると次式になり、

$$h_k(n) = h_1(n) + h_2(n) = C \cdot h_0(n) \cos \left[\frac{\pi}{2T} (2k+1)n - \frac{N-1}{2} \right] + \rho \phi_k \quad (24)$$

正の周波数では $H_k(z)$ と同じ形の振幅応答を持ち負の周波数ではゼロであるフィルタを生成する。式24のインパルス応答を持つフィルタバンクを用いると、式19のインパルス応答を持つフィルタバンクから得られるサブバンド信号に対応する。分析（複素）信号と見なしてよいサブバンド信号の集合が得られる。複素値サンプリングは極座標形式で $z(n) = r(n) + j i(n) = |z(n)| \exp(j \arg(z(n)))$ と書くことができるので、分析信号は扱いやすい。しかし、座標に複素フィルタバンクを用いると、エイリアス打消し特性を失ったために ϕ_k の制約を一般化しなければならない。エイリアス打消しと式22のインパルス応答を持つ合成フィルタバンクを保証する ϕ_k の新しい制約は

$$\phi_k = \pi(-k)^{\frac{r}{4M}} \quad (25)$$

であって、 $M=1$ のときは式21のように簡単になる。このように選択すると、座標された部分音は $M=1$ （座標なし）のときと同じ相対位相を有する。式24と式25を結合すると次式が得られる。

$$h_k(n) = C \cdot h_0(n) \cos \left[\frac{\pi}{2T} \left((-k)^{\frac{r}{4M}} - \frac{N-1}{2} \right) + \frac{(-1)^r}{4M} \right] \quad (26)$$

これは本発明のステップ4の修正されたフィルタバンクに用いるフィルタである

ステップ5について少し説明する。係数 T/M で複素値のサブバンド信号をダウンサンプリングすると M だけオーバーサンプリングされる。これは、後で位相角に座標係数 M を掛けるときの重要な判定基準である。オーバーサンプリングに

より、ターゲット範囲に座換した後の帯域幅当たりのサブバンドサンプリングの数は

原始範囲の数に等しくなる。座換されたサブバンド信号の個々の帯域幅は、位相乗算器のために原始範囲内の帯域幅の M 倍になる。このため、ステップ5の後でサブバンド信号は厳密にサンプリングされ、更に、音信号を座換するときスベクトル内にゼロがなくなる。

三角法計算を避けるために、すなわち新しいサブバンド信号を次式

$$\begin{aligned} x_k^{(M)}(n) &= \cos \left[\frac{\pi}{2} (n) \right] \cos \left[M \arctan \left(\frac{\operatorname{Im} x_k^{(M)}(n)}{\operatorname{Re} x_k^{(M)}(n)} \right) \right] = \\ &= \left| x_k^{(M)}(n) \right| \cos \left[M \arctan \left(\frac{\operatorname{Im} x_k^{(M)}(n)}{\operatorname{Re} x_k^{(M)}(n)} \right) \right]. \end{aligned} \quad (27)$$

ただし、 $|x_k^{(M)}(n)|$ は $x_k^{(M)}(n)$ の絶対値、で計算しなければならないので、次の三角法関係を用いる。

$$\cos(M\alpha) = \cos^M(\alpha) - \binom{M}{2} \sin^2(\alpha) \cos^{M-2}(\alpha) + \binom{M}{4} \sin^4(\alpha) \cos^{M-4}(\alpha) - \dots \quad (28)$$

ここで

$$\alpha = \arctan \left(\frac{\operatorname{Im} x_k^{(M)}(n)}{\operatorname{Re} x_k^{(M)}(n)} \right). \quad (29)$$

とし、また

$$\cos(\alpha) = \cos \left(\arctan \left(\frac{\operatorname{Im} x_k^{(M)}(n)}{\operatorname{Re} x_k^{(M)}(n)} \right) \right) = \frac{\operatorname{Re} x_k^{(M)}(n)}{|x_k^{(M)}(n)|}. \quad (30)$$

および

$$\sin(\alpha) = \sin \left(\arctan \left(\frac{\operatorname{Im} x_k^{(M)}(n)}{\operatorname{Re} x_k^{(M)}(n)} \right) \right) = \frac{\operatorname{Im} x_k^{(M)}(n)}{|x_k^{(M)}(n)|}. \quad (31)$$

であって、ステップ5の計算を三角法計算によらずに行うことができるので、計算は簡単になる。

M が偶数のときに座換を用いると、低域プロトタイプフィルタ $p_0(n)$ の特

性によって位相乗算器に障害が生じることがある。全ての適用可能なプロトタイプフィルタは、 z 平面内の単位円上にゼロを有する。単位円上のゼロはフィルタの位相応答を 180° シフトする。 M が偶数のとき、位相乗算器はこのシフト

を360°シフト(すなわち位相シフトが消える)と解釈する。このように位相シフトが消える周波数に位置する部分音は合成信号にエイリアシングを起こす。最悪の場合は、部分音が分析フィルタの第1サイドロープの頂点に対応する周波数の点にあるときである。振幅応答でこのロープの拒否に従って、エイリアシングの間こえかたが変わる。一例として、ISO/MPEG層1および2標準に用いるプロトタイプフィルタの第1サイドロープは96dB拒否されるが、ISO/MPEG層3標準のMDCT方式に用いるサイン窓の第1サイドロープでは拒否は23dBに過ぎない。サイン窓を用いると、この種のエイリアシングが間こえることは明らかである。この問題の解決を以下に示す。これを相対的位相同期と呼ぶ。

フィルタ $h_p(n)$ は全て連続位相応答を有する。位相角 ϕ_p は隣接チャンネルの間に相対的位相差を生じさせ、単位円上のゼロはチャンネル間で異なる周波数の位置に180°位相シフトを起こす。位相乗算器を活動化する前に隣接サブバンド信号の間の位相差を監視すれば、位相反転情報を含むチャンネルを抽出するのは容易である。音信号の場合は式25から、位相差は非反転信号では約 $\pi/2$ Mであり、したがってどちらからの信号が反転している信号では約 $\pi(1-1/2M)$ である。反転信号の検出は、隣接サブバンド内のサンプルの点乗積

$$y_p^{(M)}(n') \cdot y_{p+1}^{(M)}(n') = \text{real}\{y_p^{(M)}(n')\} \text{real}\{y_{p+1}^{(M)}(n')\} + \text{imag}\{y_p^{(M)}(n')\} \text{imag}\{y_{p+1}^{(M)}(n')\}. \quad (32)$$

を計算することにより簡単に行うことができる。式32の積が負の場合は位相差は90°より大きく、位相反転条件が存在する。ステップ5に促って複素値サブバンド信号の位相角にMを掛けると、最後に反転と印した信号は打ち消される。このように相対位相同期法を用いると180°シフトしたサブバンド信号は位相掛け算を行った後このシフトを保持して、エイリアシング打消し特性を維持する。

サブバンド包絡線調整
Stockham [「自動利得制御への一般化された線形性の適用(Application of Generalized Linearity to Automatic Gain Control)」, T. G. Stockham, Jr.,

IEEE Trans. on Audio and Electroacoustics, Vol. AU-16, No. 2, June 1968] および式1によると、音声や音楽など殆どの音は緩やかに変化する包絡線と急速に変化する一定振幅のキャリアの積で特徴づけられる。

分割帯域知覚オーディオコーダでは、オーディオ信号をフレームに区切り、サブバンドフィルタすなわち時間周波数領域変換を用いて多数の周波数帯域に分割する。殆どの型のコーデックでは、伝送または記憶のために信号をその後2つの主な信号成分であるサブバンド包絡線表現と基準化サブバンドサンプルまたは係数に分離する。以下の説明を通して「サブバンドサンプル」または「係数」とは、サブバンドフィルタから得られるサンプル値と、時間周波数変換から得られる係数を言う。「サブバンド包絡線」または「換算係数」は時間フレームにおけるサブバンドの値(各サブバンド内の平均または最大振幅など)を表し、サブバンドサンプルの基準化に用いる。しかし、サブバンド包絡線は通常原則上PCを用いて得ることもできる[米国特許番号5,684,920号]。一般的なコーデックでは、基準化されたサブバンドサンプルは、緩やかに変化する時間包絡線、したがって非常に低いビットレート(利用可能なビットレートの約10%)を用いる)で符号化されるサブバンド包絡線、に比べて高いビットレート(利用可能なビットレートの約90%)を用いる)で符号化する必要がある。

最初の信号の音色の質を保存する場合は、複製された帯域幅の正確なサブバンド包絡線が重要である。装飾または音声の知覚される音色は主に聴覚の最高オクターブにある周波数 f_{lim} より低いサブバンド分布で決まる。したがって f_{lim} より高いサブバンドの群細は余り重要でない。上記の置換法により得られる高帯域の微細構造は調整する必要がない。しかし相構造は一般に調整する必要がある。この調整を行うには、信号のサブバンド表現を濃密して包絡線の粗構造と微細構造を分離するとよい。

本発明のSBR-1実現では、高帯域の短いサブバンド包絡線はデコーダで利用可能な低帯域情報から推定する。この推定は、低帯域の包絡線を絶えず監視して、特定の規則に従って高帯域サブバンド包絡線を調整することにより行う。包絡線推定を行う新規な方法は、対象の周波数帯域空間内で漸近線を用いる。これ

は特許空間内で種々の次数の多項式により曲線の当てはめを行うことに相当する。低帯域サブバンドの上部のレベルと傾斜を推定し、この推定を用いて新しい高帯域包絡線を表す1つまたはいくつものセグメントのレベルと傾斜を定義する。最近接の交差点は周波数で固定され、ピコット点の収束をする。しかし必ずしも必要ではないが、制約を設けて高帯域包絡線軌跡を現実的な境界内に収めるのは有益である。サブバンド包絡線を推定する別の方法は、多数の代表的サブバンド包絡線のベクトル量子化VQを用いて、これをルックアップテーブルまたはコードブックに配置することである。ベクトル量子化は大量の訓練データ（この場合はオーディオサブバンド包絡線）上の所望の数のベクトルを訓練することにより行う。この訓練は通常は一般化されたロイド(Lloyd)アルゴリズムで行い〔「ベクトル量子化と信号圧縮(Vector Quantization and Signal Compression)」, A. Gersho, R. M. Gray, Kluwer Academic Publishers, USA 1992, ISBN 0-7923-9181-0〕、訓練データの内容を最適にカバーするベクトルを生成する。B包絡線で訓練されたAサブバンド包絡線(B>A)から成るVQコードブックを考えると、A包絡線は、多くの種類の音のB個の傾斜に基づく、低帯域包絡線から高帯域包絡線へのA個の最も可能性のある遷移を表す。これは理論的には、B個の傾斜に基づいて包絡線を予測するためのA個の最適規則である。新しい高帯域サブバンド包絡線を推定するときは、初めの低帯域包絡線を用いてコードブックを探し、コードブックの最も合致する項目の高帯域域を用いて新しい高帯域サブバンド包絡線を作成する。

図23に、基調化されたサブバンドサンプルを2301で表し、サブバンド包絡線を換算係数2305で表す。例示のために、デコード2303への伝送を並列形式で示す。SBR-2法の図24では、図23と同様にサブバンド包絡線情報を生成して伝送するが、サブバンドサンプルは低帯域だけを伝送する。したがって伝送される換算係数は全周波数範囲にわたるが、サブバンドサンプルは高帯域を除く限られた周波数範囲だけである。デコードで低帯域サブバンドサンプル2401を2403のように遷移し、受信した高帯域サブバンド包絡線情報2405と結合する。このようにすれば、合成高帯域サブバンド包絡線は初めのサブバンド包絡線と同じであるがビットレートは大幅に下がる。

或るコーデックでは、図24に示すように全サブバンド包絡線の換算係数を伝送し、高帯域サブバンドサンプルは省略することができる。他のコーデック形式では、換算係数とサブバンドサンプルが同じ周波数範囲をカバーするよう規定しなければならない。すなわち、サブバンドサンプルを省略した場合は換算係数を伝送することはできない。この場合はいくつかの解決法がある。例えば、高帯域サブバンド包絡線情報を別のフレームで伝送してよい。このフレームは、自身のヘッドと任意の誤り保護を持ち、その後にデータが続く。本発明を利用しない普通のデコーダはヘッドを認識しないので、余分なフレームは廃棄する。第2の解決法では高帯域サブバンド包絡線情報を、符号化されたビットストリーム内の補助データとして伝送する。しかし利用可能な補助データフィールドは包絡線情報を保持するだけ十分大きくない。この2つの解決法が適用できない場合は、第3の解決法、すなわち高帯域サブバンド包絡線情報をサブバンドサンプルと異なる大きな動的範囲をカバーする。任意の数のサブバンド換算係数(図25の2505)を非常に低い値に設定して、高帯域換算係数をサブバンドサンプルとして「偽装し」て2501のように伝送することができる。このように高帯域換算係数をデコード2503に伝送することにより、ビットストリーム構文と両立させることができる。任意のデータをこの方法で伝送してよい。これに関連して、情報を符号化してサブバンドサンプルストリームにする方法がある〔米国特許番号5,687,191号〕。図26に示す第4の解決法は、符号化システムがハフマンまたは他の冗長度符号化2603を用いるときに適用することができ、高い冗長度を達成するには、高帯域のサブバンドサンプルをゼロ(2601)にまたは一定値に設定する。

過渡伝送の改善

過渡信号に関連する人工音はオーディオコーデックの共通の問題であり、同様な人工音は本発明でも発生する。一般に、パッチングを行うと時間領域の前エコーと後エコー（すなわち「真の」過渡信号の前か後の疑似過渡信号）に相当するサブバンド「ゼロ」すなわちノッチを生成する。Pブロックはゆっくり変化する音信号の「ゼロを埋める」が、前エコーと後エコーは残る。改善された多帯域法

は、正弦波の数がサブバンド当たり1個に制限された離散的な正弦波に作用するものである。サブバンド内の過渡信号すなわち雑音は、そのサブバンド内の多数の離散的な正弦波と見ることができ、これは相互変調歪みを生成する。これらの人工音は、複製された高帯域チャンネルに過渡期間中に登録された追加の量子化雑音源と考えられる。したがって、知覚オーディオコーデック内の前エコーおよび後エコー人工音を避ける従来の方法（例えば適応的エコー置換）を用い、改善された多帯域法的主観的品質を高めることができる。コーデックまたは別個の抽出器による過渡検出を用い、また過渡状態にあるチャンネル数を減らせば、「量子化雑音」は時間的に放散するマスキングしきい値を超えない。過渡信号の通過中は少数のチャンネルを用い、音の通過中は多数のチャンネルを用いる。このような適応的なエコー置換はコーデック内で普通に用いられており、周波数分解能と時間分解能の間で取引する。フィルタバンク大きさが固定されている応用には別の方法を用いてよい。1つの方法は、サブバンド領域内の連続予測により「量子化雑音」を時間で成形することである。次に残留信号に置換を行う。これが連続予測フィルタの出力である。その後反転予測フィルタを、切りのチャンネルとサブバンド複製チャンネルに同時にかける。別の方法はコンパンダシステム（すなわち、置換または符号化の前の過渡信号の動的振幅圧縮と、置換の後の確定的伸張）を用いる。また、信号に依存して置換法の間で切り替えることもできる。例えば、定率信号に高分解能フィルタバンク置換法を用い、過渡信号に時域パターン探索予測法を用いる。

実験的な応用

標榜の信号プロセッサまたは強力PFCを用いると、SBR強化コーデックを長時間で動作させることができる。SBR強化コーデックはカスタムチップにハードで符号化してもよい。また図27や図28のように任意のコーデックを用いて種々のシステムでこれを実現して、アナログ信号またはディジタル信号の配信または伝送に用いてよい。SBR-1法は、デコードに組み込んでも、付加的なハードウェアまたはソフトウェア後処理モジュールとして供給してもよい。SBR-2法はエンコーダを更に修正する必要がある。図27において、アナログ入力信号がA/D変換器2701に入り、ディジタル信号を形成して任意のエンコー

ダ2703に与え、ここで原始コーデッキングを行う。この装置に入る信号は、聴覚範囲内のサブバンド帯域をすでに廃棄した。またはサブバンド帯域を任意のエンコーダ内で廃棄した低域信号でよい。得られる低帯域信号をマルチプレクサ2705に与えて直列ビットストリームを形成し、2707で伝送または配信する。マルチプレクサ2709は信号を回復して任意のデコード2711に与える。サブバンド包絡線情報2715をデコード2711で評価してSBR-1ユニット2713に与える。ユニット2713は低帯域信号を高帯域信号に置換して、包絡線を調整した広帯域信号を生成する。最後に、2717でディジタル広帯域信号をアナログ出力信号に変換する。

SBR-2法はエンコーダを更に修正する必要がある。図28において、アナログ入力信号がA/D変換器2801に入り、ディジタル信号を形成して任意のエンコーダ2803に与え、ここで原始コーデッキングを行う。2805でサブバンド包絡線情報を取り出す。得られる信号は低帯域サブバンドサンプルまたは係数と広帯域包絡線情報で、これをマルチプレクサ2807に与えて直列ビットストリームを形成し、2809で伝送または配信する。マルチプレクサ2811は信号や低帯域サブバンドサンプルまたは係数や広帯域包絡線情報を回復して、任意のデコード2815に与える。サブバンド包絡線情報2813をマルチプレクサ2811からSBR-2ユニット2817に送り、低帯域信号を高帯域信号に置換して、包絡線を調整した広帯域信号を生成する。最後に、2819でディジタル広帯域信号をアナログ出力信号に変換する。

非常に低いビットレートだけしか利用できないときは（インターネットや、悪い電話モデム、AM放送など）、オーディオプロログラム材料のモノコーデッキングが避けられない。知覚品質を高め、より快適な音をプログラムするには、タッピング遅延線2901を導入すれば図29に示す簡単な「疑似ステレオ」発生器が得られる。これは、10msと15msの遅延信号を2903で約-6dBにして各出力チャンネルに与え、2905で切りのモノ信号に加える。疑似ステレオ発生器を用いると、低い計算コストで大きな知覚改善が得られる。

上述の装置の形態は、オーディオ原始コーデッキングを改善するための本発明の原理を示すものに過ぎない。理解されるように、ここに述べた装置や詳細の修正

や変更は当業者には明らかである。したがって、課せられる制約は特許請求の範囲だけによるものであって、ここで実施の形態の記述や説明により示した特定の詳細によるものではない。

【図1】

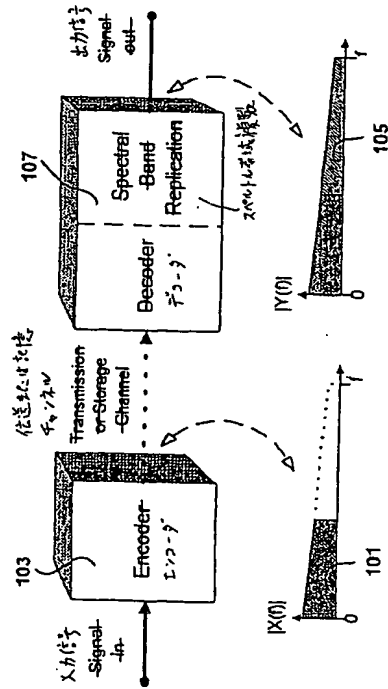


Fig. 1

【図2】

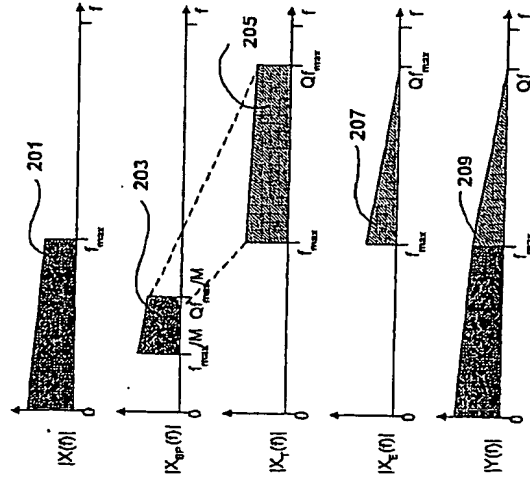


Fig. 2

[図 4]

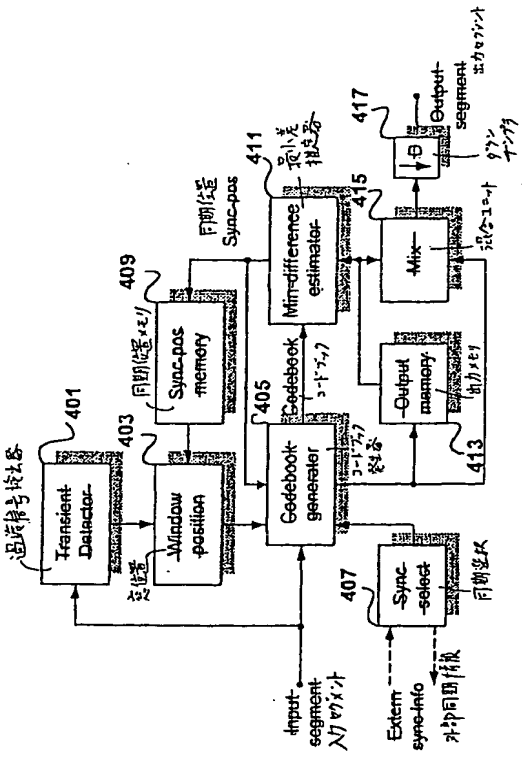


Fig. 4

[図 3]

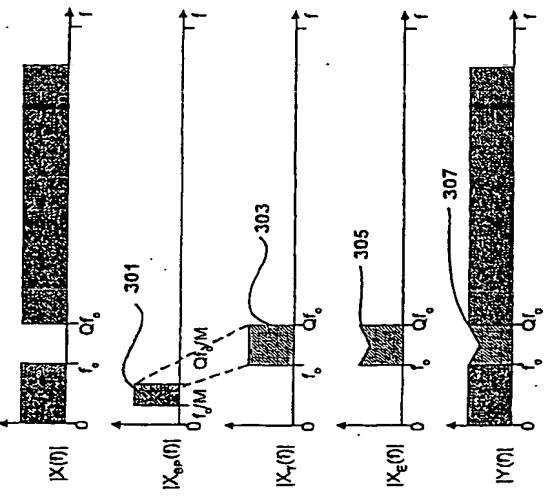


Fig. 3

【図5】

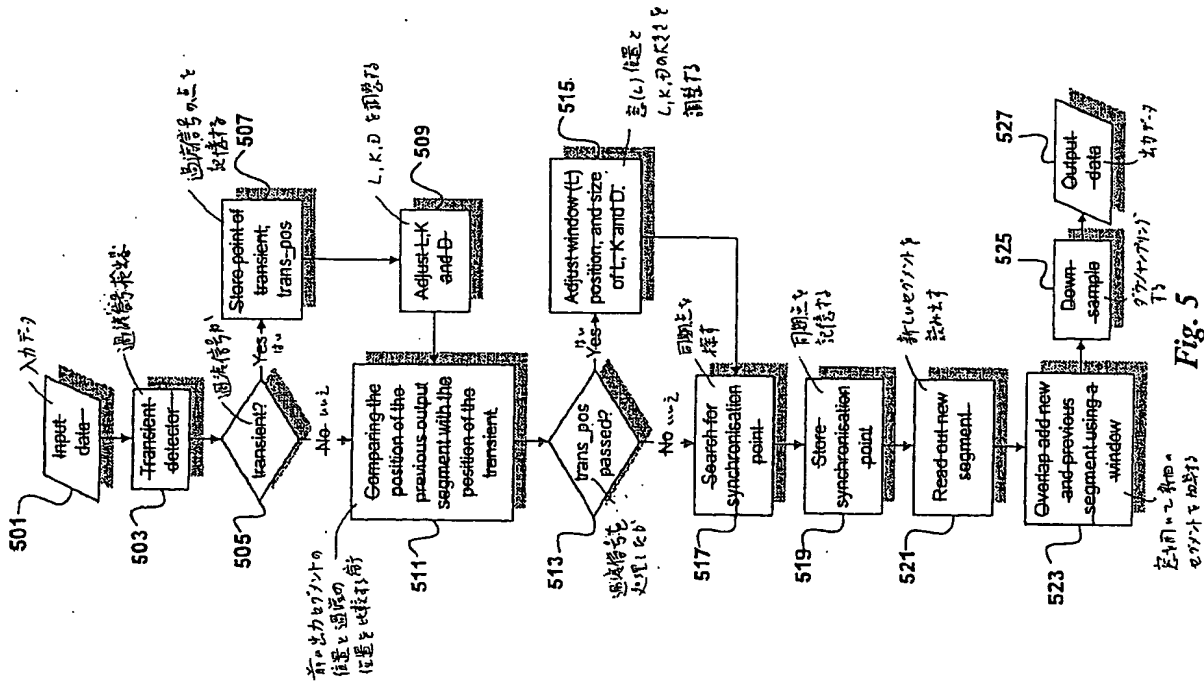


Fig. 5

【図6】

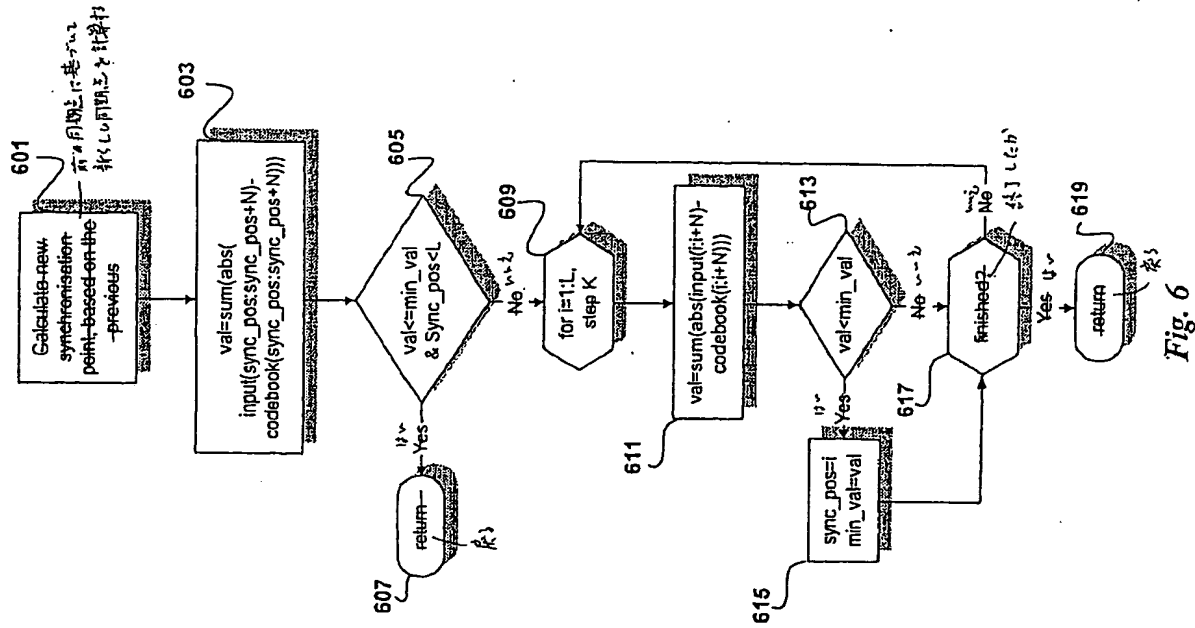


Fig. 6

【図8】

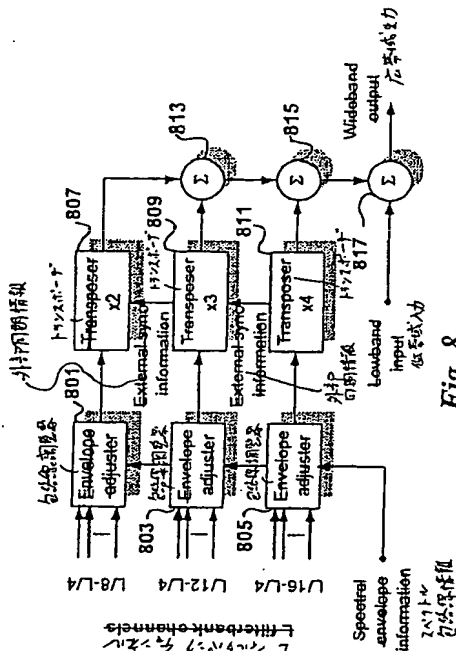


Fig. 8

【図7】

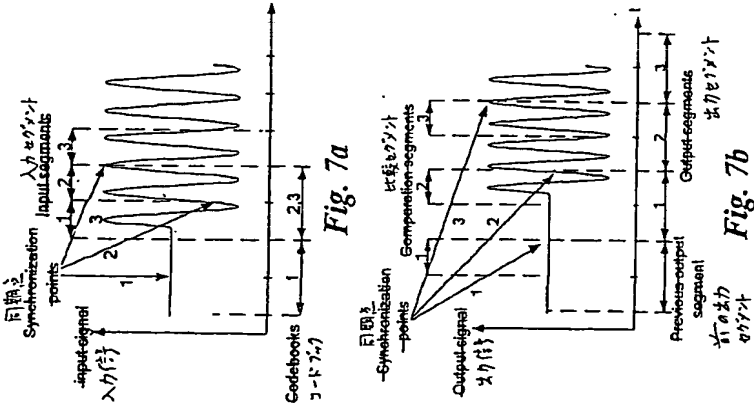


Fig. 7a

Fig. 7b

[图9]

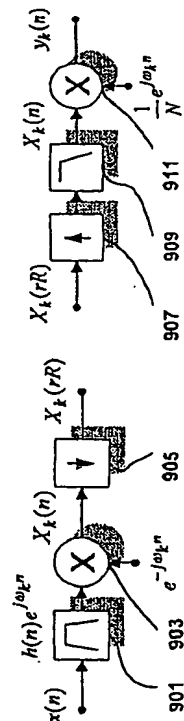


Fig. 9a

Fig. 9b

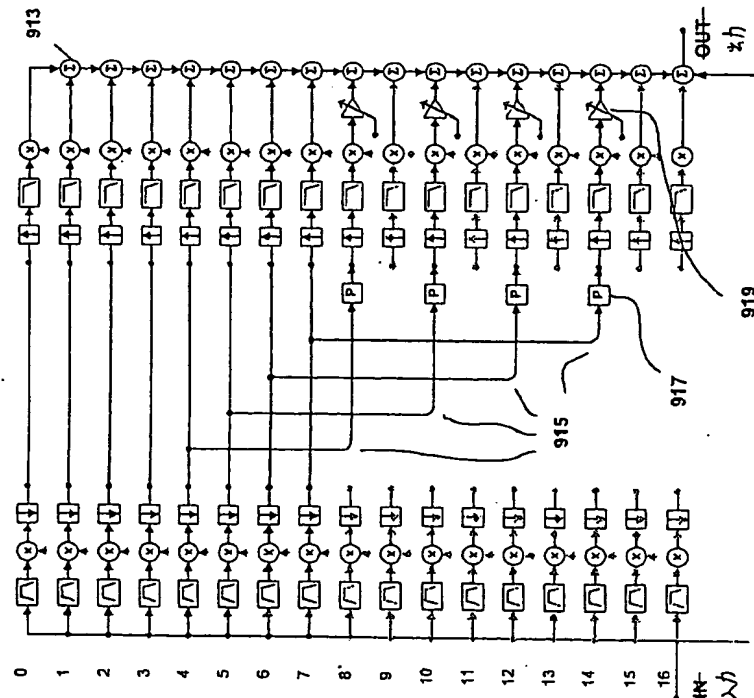


Fig. 9c

[图10]

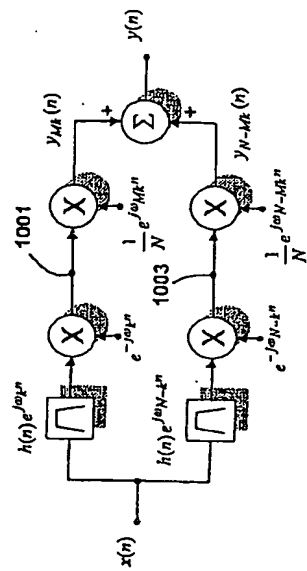


Fig. 10a

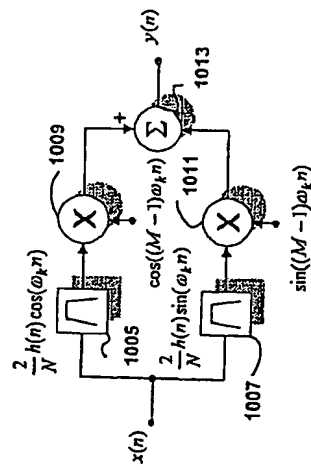


Fig. 10b

[图11]

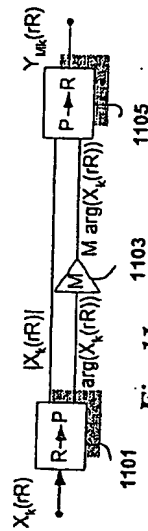


Fig. 11

[圖 12]

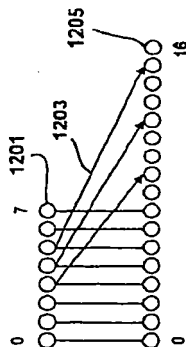


Fig. 12

[圖 13]

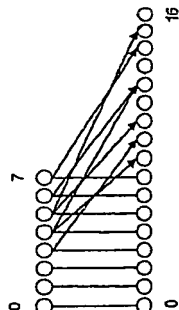


Fig. 13

[圖 14]

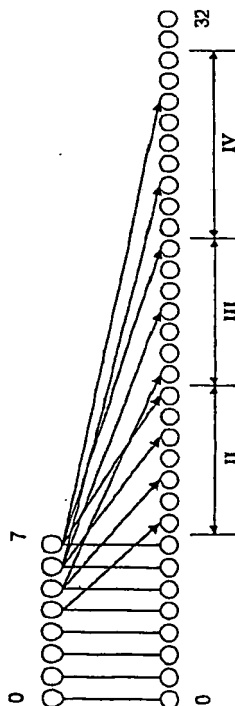


Fig. 14

[圖 15]

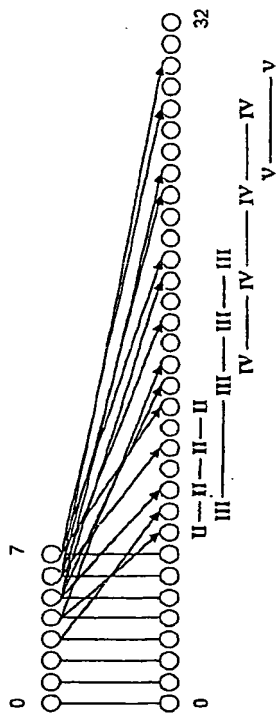


Fig. 15

[圖 16]

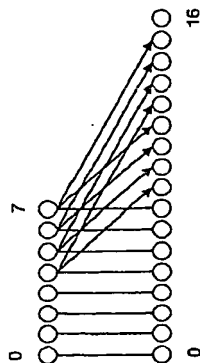


Fig. 16

[圖 17]

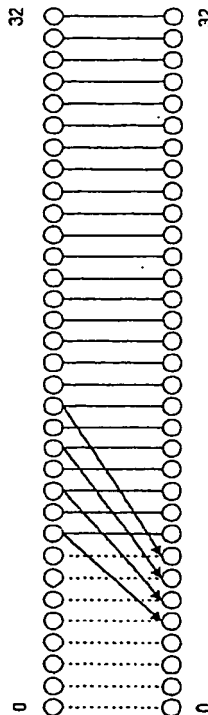


Fig. 17

[図18]

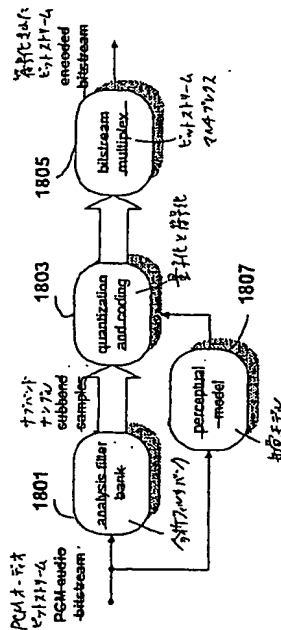


Fig. 18a

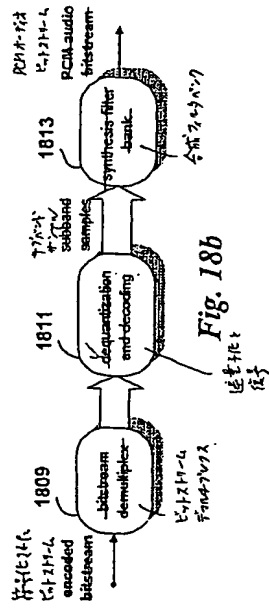


Fig. 18b

[図19]

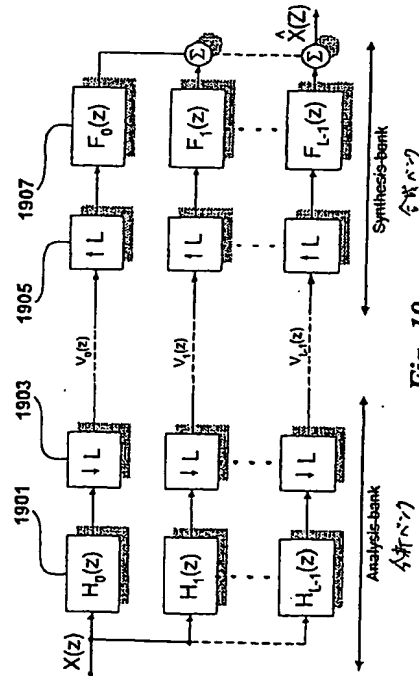


Fig. 19

[図20]

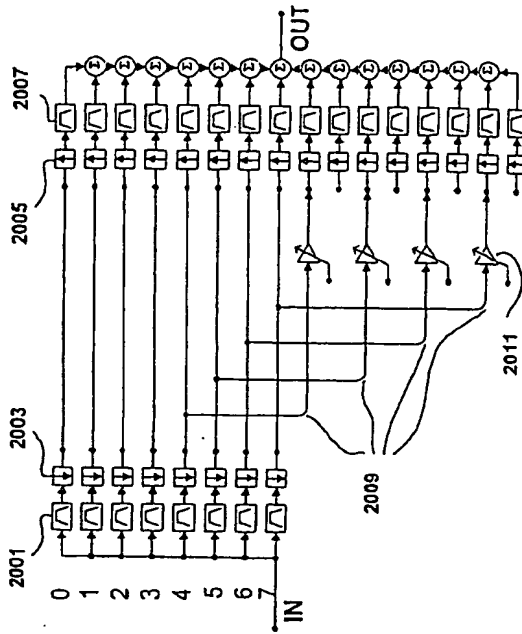


Fig. 20

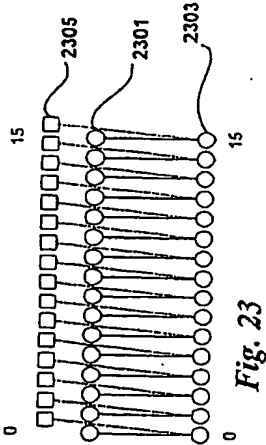


Fig. 23

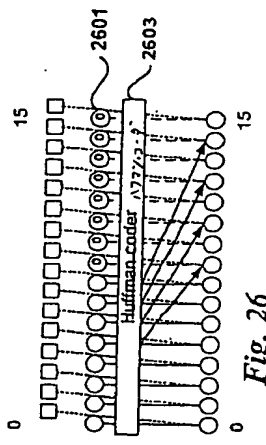


Fig. 26

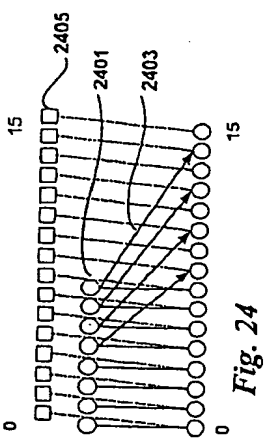


Fig. 24

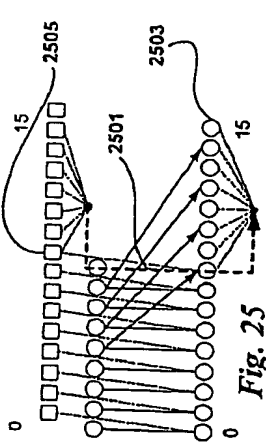


Fig. 25

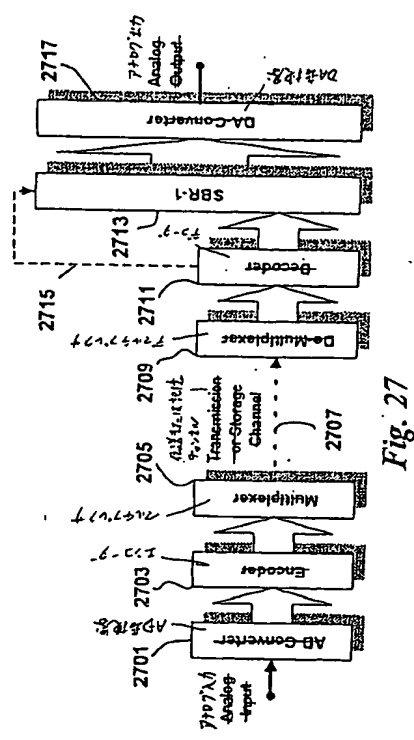


Fig. 27

[図 28]

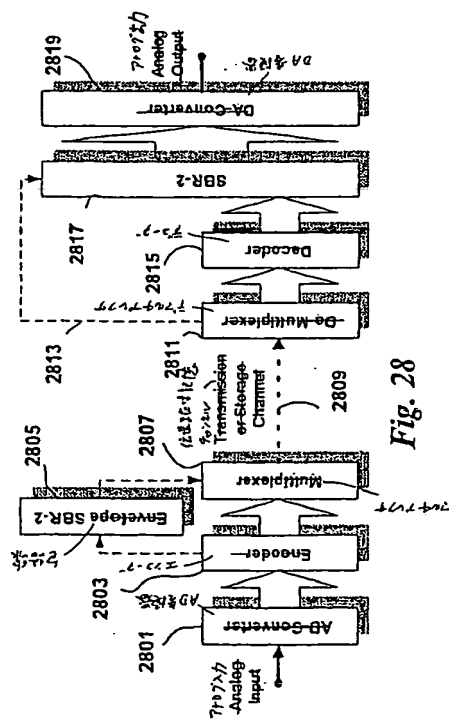


Fig. 28

[図 29]

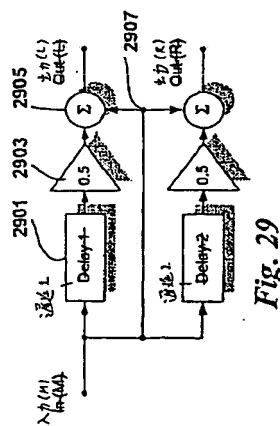


Fig. 29

【手続補正書】

【提出日】平成13年6月11日（2001. 6. 11）

【補正内容】

請求の範囲

1. 原始コーデイング装置の強化方法であって、前記原始コーデイング装置は、配送または伝送の前に行う全ての操作を表すエンコードと、配送または伝送の後に行う全ての操作を表すデコードを含み、前記エンコーダで初めの信号の周波数帯域を被棄して第1信号を形成し、前記デコーダで前記第1信号に置換を行って前記初めの信号の周波数帯域を複製して第2信号を形成し、前記第1と第2信号を結合して出力信号を形成し、これにより所定の知覚品質でビットレートを下げ、または所定のビットレートで知覚品質を高める、ことを特徴とする、原始コーデイング装置の強化方法、
2. 前記第2信号の通過帯域は前記第1信号の通過帯域と重ならないまたは一掃だけ重なるように設定することを特徴とする、請求項1に記載の原始コーデイング装置の強化方法、
3. スペクトル包絡線調整は、前記第1信号を用いて、前記初めの信号の前記被棄された周波数帯域のスペクトル包絡線の推定に基づいて行うことを特徴とする、請求項1～2のいずれかに記載の原始コーデイング装置の強化方法、
4. スペクトル包絡線調整は、前記初めの信号の前記被棄された周波数帯域の、伝送された包絡線情報に基づいて行うことを特徴とする、請求項1～2のいずれかに記載の原始コーデイング装置の強化方法、
5. 前記スペクトル包絡線情報は、その利得が低レベルに設定された任意の数のサブバンドチャネル内のサブバンドサンプルとして伝送され、これにより増強されたデコードとの互換性が確保されることを特徴とする、請求項4に記載の原始コーデイング装置の強化方法、
6. 前記包絡線情報を換算係数として伝送し、対応するサブバンドサンプルは伝送しないことを特徴とする、請求項4に記載の原始コーデイング装置の強化方法、

7. 前記包絡線情報を換算係数として伝送し、前記対応するサブバンドサン
プルをゼロまたは一定値に設定することにより、前記サブバンドサンプリングのエン
トロピーを減らすことを特徴とする、請求項4に記載の原始コーディング

装置の強化方法。

8. モノフォニックオーディオのときは、前記出力信号を、前記出力信号と
それを遅延した信号をそれぞれ含む2つの信号に分割して疑似ステレオ信号を得
ることを特徴とする、請求項1-7のいずれかに記載の原始コーディング装置の
強化方法。

9. 前記置換は、
信号を、それぞれ周波数 $\{f_1, \dots, f_N\}$ を含む通過帯域を持つN個 (N
≥ 2) の帯域フィルタの集合で濾波して、N個の帯域信号を形成し、
前記帯域信号の周波数を、周波数 $\{f_1, \dots, f_N\}$ を含む領域にシフト
し (ただし、 $M \neq 1$ は置換係数)。

前記シフトされた帯域信号を結合して置換された信号を形成する、
ことを特徴とする、請求項1-7のいずれかに記載の原始コーディング装置の強
化方法。

10. 前記周波数シフトを上側帯域 (USB) 変調により得ることを特徴と
する、請求項9に記載の原始コーディング装置の強化方法。

11. 係数Mで置換する方法であって、
信号を、低帯域型の実数値または複素値のサブバンド信号を生成する性質の分
析フィルタバンクまたは変換を用いて帯域濾波し、

合成フィルタバンクまたは変換内で、前記分析フィルタバンクまたは変換の任
意の数のチャンネルをチャンネルMk ($M \neq 1$) にパッチングし、

前記合成フィルタバンクまたは変換を用いて、置換された信号を形成する、こ
とを特徴とする、係数Mで置換する方法。

12. 前記フィルタバンクを最大10進化し、前記パッチングを次の関係によ
り行い、

$$v_{kL}(n) = (-1)^{(M-1)kn} v_k(n),$$

ただし、 $(-1)^{(M-1)kn}$ は訂正係数、 $v_k(n)$ はチャンネルkのサブバンド信
号、 $v_{kL}(n)$ はチャンネルMkのサブバンド信号であり、これによりスペクト
ル反転サブバンド信号の補正が得られることを特徴とする、請求項1

1に記載の係数Mで置換する方法。

13. 前記分析フィルタバンクまたは変換のチャンネルkからのサブバンド
信号の位相を、合成チャンネルMk ($M \neq 1$) に関連するサブバンド信号の位相
としてパッチングし、

前記分析フィルタバンクまたは変換の連続的なチャンネル1からのサブバンド
信号の振幅を、連続的な合成チャンネル1+S (Sは整数≠1) に関連するサブ
バンド信号の振幅としてパッチングする、

ことを特徴とする、請求項1-12のいずれかに記載の係数Mで置換する方法

14. 前記合成フィルタバンクまたは変換を用いる前に、前記チャンネルk
の前記サブバンド信号の位相に前記係数Mを掛けることを特徴とする、請求項1
1-13のいずれかに記載の係数Mで置換する方法。

15. $M-K \neq 1$ (ただし、Kは整数>1)であることを特徴とする、請求
項11-14のいずれかに記載の係数Mで置換する方法。

16. 前記パッチングは前記置換係数Mの多量の値を用いることを特徴とす
る、請求項11-15のいずれかに記載の係数Mで置換する方法。

17. 係数Mで置換する方法であって、
インパルス応答

$$h_k(n) = K \rho(n) \exp \left[j \frac{\pi}{2L} (2k+n)(n - \frac{M-1}{2}) + j(n)^2 \frac{\pi}{4M} \right],$$

ただし、 $k=0, 1, \dots, L-1$, Kは定数、 $\rho_0(n)$ は長さNの低域プ
ロトタイプフィルタ、を持つL個のフィルタの並列バンクで信号を濾波して、L
個の複素値信号の集合を生成し、

係数 L/M を持つ前記 L 個の信号をダウンサンプリングして、 L 個の複素値サブバンド信号の集合を生成し、
前記複素値サブバンド信号の位相角に M を掛けて、サブバンド信号の新しい集合を生成し、
前記サブバンド信号の新しい集合の実数部を選択して、 L 個の実数値サブバンド信号の集合を生成し、

係数 L' を持つ前記実数値サブバンド信号の部分集合をアップサンプリングして、実数値信号の集合を生成し、
インパルス応答

$$h(n) = K' p_0(n) \cos \left[\frac{\pi}{2L'} (2k + 1) \left(\alpha - \frac{M' - 1}{2} \right) - (-1)^k \frac{\pi}{4} \right].$$

ただし、 $k = 0, 1, \dots, L' - 1$ 、 K' は定数、 $p'_0(n)$ は長さ N' の低域プロトタイプフィルタを持つ L' 個のフィルタの並列バンクで前記実数値信号を濾波して、 L' 個の濾波信号の集合を形成し、
前記 L' 個の濾波信号を加算して復調信号を生成する、
ことを特徴とする、係数 M で復調する方法、
18. 前記位相角の前記掛け算と前記実数部の前記選択を計算するのに、前記複素値サブバンド信号を次式で書き、

$$Z_k(n) = R_k(n) + jI_k(n).$$

ただし、 $R_k(n)$ と $I_k(n)$ はそれぞれ $Z_k(n)$ の実数部と虚数部であり、
前記実数値サブバンド信号 $w_k(n)$ を次式で計算し、

$$w_k(n) = |Z_k(n)| \cos \left\{ M \arctan \left(\frac{I_k(n)}{R_k(n)} \right) \right\}.$$

ただし、 $|Z_k(n)| = \sqrt{R_k(n)^2 + I_k(n)^2}$ 、 M は正の整数の置換係数であり、次の三角恒等式

$$\cos(M\alpha) = \cos^M(\alpha) - \binom{M}{2} \sin^2(\alpha) \cos^{M-2}(\alpha) + \binom{M}{4} \sin^4(\alpha) \cos^{M-4}(\alpha) - \dots,$$

ただし、 $\alpha = \arctan(I_k(n)/R_k(n))$ 、と次の関係

$$\cos(\alpha) = \frac{R_k(n)}{|Z_k(n)|} \text{ and } \sin(\alpha) = \frac{I_k(n)}{|Z_k(n)|}$$

を用い、これにより全ての三角法計算をなくして計算の複雑さを減らす、ことを特徴とする請求項17に記載の係数 M で復調する方法、
19. ブロック毎に、前記複素値サブバンド信号の隣接対の位相差により選ばれる情報を取り出し、

前記位相角に前記 M を掛けて前記新しいサブバンド信号の対を形成し、
前記情報を与える条件によって前記新しいサブバンド信号の1つを打ち消すことにより、偶数整数値の置換係数 M を用いるときにサブバンド信号の 180° 位相シフトを保持する、

ことを特徴とする、請求項17に記載の係数 M で復調する方法、

20. 前記情報は次式の前記複素値サブバンド信号 $Z_k(n)$ と $Z_{k+1}(n)$ の点乗積で与えられ、

$$Z_i(n) \cdot Z_{i+1}(n) = R_i(n)R_{i+1}(n) + jI_i(n)I_{i+1}(n),$$

ただし、 $R_i(n)$ と $I_i(n)$ はそれぞれ $Z_i(n)$ の実数部と虚数部 ($i = k, k+1$) であり、前記点乗積が負の場合は前記新しいサブバンド信号の1つを打ち消すことを特徴とする、請求項19に記載の係数 M で復調する方法、

21. 第1信号を時間的に伸張または圧縮し、前記第1信号の任意の長さのセグメントを複写または廃棄し、次に前記第1信号をダウンサンプリングまたはアップサンプリングする、置換方法であって、
前記第1信号に過渡検出を行い、
過渡検出の結果に従って、前記第1信号の一部を複写または廃棄するときに前記第1信号のどのセグメントを用いるかを決定し、
前記過渡検出の結果に従って前記信号セグメントの長さ L を調整し、
前記過渡検出の結果に従って各状態ベクトルに用いるサンプル数 N を調整し、
前記過渡検出の結果に従って状態ベクトル内のサンプル間の遅れ D を調整し、
前記過渡検出の結果に従って各状態ベクトル間のサンプル数 K を調整し、
前の同相点探索で見出した同相点に基づいて、前記第1信号の選択されたセグ

メント内の同期点を検ず、
ことを特徴とする置換方法。

22. いくつかのトランスポートを相互接続して同期点情報を共有して、計算の複雑さを減らすことを特徴とする、請求項21に記載の置換方法。

23. 前記トランスポートを適当なフィルタバンクに接続し、前記各トランスポートに与える信号を増幅して、前記トランスポートが処理中の前記信号

の和である新しい信号の任意のスペクトル包絡線を得ることを特徴とする、請求項21-22のいずれかに記載の置換方法。

24. 切めの信号から得られる原始コーディング信号の復号を強化する装置であって、

前記原始コーディング信号の周波数帯域を置換して第1信号を形成する置換手段と、

前記原始コーディング信号に作用して前記切めの信号のスペクトル包絡線を推定する推定手段と、

前記推定に基づいて、前記第1信号のスペクトル包絡線を調整する調整手段と

前記原始コーディング信号と前記調整された第1信号を結合して、所定の知覚品質でビットレートを下げ、または所定のビットレートで知覚品質を高める、結合手段、
を特徴とする、原始コーディング信号の復号を強化する装置。

25. 前記出力信号がモノフォニックオーディオのときに動作し、

第1遅延信号を形成するための、前記出力信号を選らせる遅延手段および減衰させる減衰手段と、

第2遅延信号を形成するための、異なるパラメータを用いる、前記出力信号を選らせる遅延手段および減衰させる減衰手段と、

前記出力と前記第1遅延信号を加算して左チャンネル出力信号を形成する手段と、

前記出力と前記第2遅延信号を加算して右チャンネル出力信号を形成して、疑

似ステレオフォニック信号を得る手段、

を特徴とする、請求項24に記載の原始コーディング信号の復号を強化する装置

26. 原始コーディングの強化装置であって、前記装置は配信媒体または伝送チャンネルの前の全てのユニットを表すエンコードと、前記配信媒体または伝送チャンネルの後の全てのユニットを表すデコードを含むものであり、その特徴は、

前記エンコードで切めの信号の周波数帯域を置換して第1信号を形成する置換手段と、

前記エンコードで前記切めの信号のスペクトル包絡線情報を取り出して第2信号を形成する取出し手段と、

前記エンコードで前記第1信号と第2信号を符号化する手段と、

前記デコードで前記第1信号の周波数帯域を置換して第3信号を形成する置換手段と、

前記第2信号に基づいて、前記デコードで前記第3信号のスペクトル包絡線を調整する調整手段と、

前記デコードで前記第1信号と前記調整された第3信号を結合して、所定の知覚品質でビットレートを下げ、または所定のビットレートで知覚品質を高める、結合手段、

である原始コーディングの強化装置。

27. 前記出力信号がモノフォニックオーディオのときに動作し、

第1遅延信号を形成するための、前記出力信号を選らせる遅延手段および減衰させる減衰手段と、

第2遅延信号を形成するための、異なるパラメータを用いる、前記出力信号を選らせる遅延手段および減衰させる減衰手段と、

前記出力と前記第1遅延信号を加算して左チャンネル出力信号を形成する手段と、

前記出力と前記第2遅延信号を加算して右チャンネル出力信号を形成して、疑

似ステレオファニック信号を得る手段、
を特徴とする。請求項26に記載の原格コーディングの強化装置。

28. 係数Mで置換する装置であって、
信号を、低帯域型の実数値または複素数値サブバンド信号を生成する性質の分析フィルタバンクまたは変換により帯域透過することと、
合成フィルタバンクまたは変換により帯域透過することと、
合成フィルタバンクまたは変換により帯域透過することと、
意のチャネル数kをチャネルMk (M≠1) にパッチングする手段と、
前記合成フィルタバンクまたは変換により、置換された信号を形成すること、
を特徴とする。

を特徴とする係数Mで置換する装置。

29. M=K±1のとき (Kは整数>1)、前記合成フィルタバンクまたは変換を用いる前に前記チャネルkのサブバンド信号の位相にMを掛けることを特徴とする。請求項28に記載の係数Mで置換する装置。

30. 係数Mで置換する装置であって、
インパルス応答

$$h_k(n) = K \rho_0(n) \exp \left[j \frac{\pi}{2L} (2k + N) \left(n - \frac{N-1}{2} \right) + j (n - \frac{N-1}{2}) \frac{\pi}{4} \right].$$

ただし、k=0, 1, ..., L-LKは定数、 $\rho_0(n)$ は長さNの低域プロトタイプフィルタ、Mは置換係数、を持つL個のフィルタの並列バンクで信号を透過して、L個の複素数値信号の集合を生成する濾波手段と、

係数L/Mを持つ前記L個の信号をダウンサンプリングして、L個の複素数値のサブバンド信号の集合を生成する手段と、

前記複素数値サブバンド信号の位相角にMを掛けて、サブバンド信号の新しい集合を生成する手段と、

前記サブバンド信号の新しい集合の実数部を選択して、L個の実数値サブバンド信号の集合を生成する手段と、

係数L'を持つ前記実数値サブバンド信号の部分集合をアップサンプリングして、実数値信号の集合を生成する手段と、
インパルス応答

$$f_k(n) = K' \rho'_0(n) \exp \left[j \frac{\pi}{2L'} (2k + N') \left(n - \frac{N'-1}{2} \right) + j (n - \frac{N'-1}{2}) \frac{\pi}{4} \right].$$

ただし、k=0, 1, ..., L'-1、K'は定数、 $\rho'_0(n)$ は長さN'の低域プロトタイプフィルタ、を持つL'個のフィルタの並列バンクで前記実数値信号を濾波して、L'個の複素数値信号の集合を形成する濾波手段と、
前記L'個の複素数値信号を加算して置換信号を生成する手段、
を特徴とする。係数Mで置換する装置。

31. 第1信号を時間的に伸張または圧縮し、前記第1信号の任意の長さのセグメントを複写または廃棄し、次に前記第1信号をダウンサンプリングま

たはアップサンプリングする、置換装置であって、

前記第1信号に過渡検出を行う検出手段と、
可能な過渡信号の位置を用いて、前記第1信号の一部を複写または廃棄すると共に前記第1信号のどのセグメントを用いるかを決定して、前記置換を得る手段と、

前記過渡検出器からの出力に従って前記信号セグメントの長さ(L)を調整する調整手段と、

前記過渡検出器からの出力に従って各状態ベクトルに用いるサンプル数(N)を調整する調整手段と、

前記過渡検出器からの出力に従って前記状態ベクトル内のサンプル間の遅れ(D)を調整する調整手段と、

前記過渡検出器からの出力に従って各状態ベクトル間のサンプル数(K)を調整する調整手段と、

前記同相点探索で見出した同相点に基づいて、前記第1信号の選択されたセグメント内の同相点を探査手段、
を特徴とする置換装置。

32. サブバンド信号に作用して、
前記トランスポーズの多数の事例の間で同相値幅を共用する手段と、
前記サブバンド信号の部分集合を形成する手段と、

前記各部分集合内で各チャネルの振幅調整を行う手段と、
前記各部分集合から、前記トランスポザの各事例への入力番号を形成する合
成ファイルバンク手段と、
前記トランスポザによる前記入力番号の処理と、
前記処理番号を加算することにより新しい番号を得て、任意のサブキャリア包絡
線を得る加算手段、
特徴とする、請求項31に記載の周波数変換装置。

【國際調查報告】

INTERNATIONAL SEARCH REPORT

9. CLASSIFICATION OF SUBJECT MATTER IPC 6 H0481/06		International Application No I / 118 98 / 00893	
INTERNATIONAL SEARCH REPORT			
According to International Patent Class (IPC) or to both national classifications and IPC			
8. FIELD OF SEARCHED			
IPC 6 H048			
Documentation concerning other than minimum documentation to be submitted in the fields searched			
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)			
C. DOCUMENTS CONSIDERED TO BE RELEVANT		Referent to claim No.	
Category *	Citation of document, with indication, where appropriate, of the relevant passages	1-4, 9, 11, 24, 26, 28	
X	US 5 127 054 A (HONG DAERYOUNG ET AL) 30 June 1992 cited in the application see figure FIGURE see column 1, line 41 - line 50 see column 1, line 65 - column 2, line 2 see column 2, line 11 - column 3, line 33 ---	1-4, 24, 26 11-14	
X	US 4 667 340 A (ARIMAND NASUD ET AL) 19 May 1997	---	
A	see abstract; figures 4, 6 see column 3, line 8 - line 43 see column 7, line 48 - column 8, line 22 ---	-/-	
* Further documents are listed in the continuation of list C.		Patent family members are listed in annex.	
* Special categories of other documents :			
"A" document defining the general state of the art which is not considered to be of particular relevance			
"B" other document but published on or after the international			
filing date which may influence directly or indirectly the			
"C" document which may influence directly or indirectly the			
"D" document which may influence the publication date of another			
"E" document referring to an amendment, variation, extension or			
"F" document published prior to the international filing date but			
later than the priority date claimed			
Date of the actual completion of the international search		Date of setting of the international search report	
26 April 1999		21 12 1999	
Name and mailing address of the ISA TM, 2210 JAV Road, Paris (93) 93 500-5046, 75 018 09 4, Fax (93) 93 500-5015		Authorized officer Sieben, S	

INTERNATIONAL SEARCH REPORT

Informational Application No
F 118 98/00893

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		Relevant to identify
Category	Chapters of document, with indication, where appropriate, of the relevant paragraph	
A	<p>US 5 068 899 A (ELLIS JOHN G ET AL) 26 November 1991 see abstract; figures 1-7 see column 1, line 39 - line 52 see column 2, line 1 - line 36 see column 2, line 64 - column 3, line 5 see column 5, line 62 - column 6, line 14</p> <p>---</p> <p>A J S FERREIRA: "Audio Spectral Coder" AES REPRINT 4201, 108TH CONVENTION, COPENHAGEN, 11 - 14 May 1996, XP002100534 cited in the application see figure 3 see page 5, paragraph 4.2; figures 2, 4, 5 see page 7, paragraph 5; figure 4</p> <p>----</p>	<p>1-4, 6, 24, 26</p> <p>1-4, 7, 17, 24, 26, 30</p>
A		

U.S. DEPARTMENT OF JUSTICE
FEDERAL BUREAU OF INVESTIGATION
WASHINGTON, D.C. 20535

INTERNATIONAL SEARCH REPORT

ational application No.
PCT/18 98/00893

<p>Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)</p> <p>The International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:</p> <ol style="list-style-type: none"> 1. <input type="checkbox"/> Claims Nos.: _____ because they relate to subject matter not required to be searched by this Authority, namely: 2. <input type="checkbox"/> Claims Nos.: _____ because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: 3. <input type="checkbox"/> Claims Nos.: _____ because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a). 	<p>Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)</p> <p>The International Searching Authority found multiple inventions in this International application, as follows:</p> <p style="text-align: center; margin: 10px 0;">see additional sheet</p> <ol style="list-style-type: none"> 1. <input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims. 2. <input type="checkbox"/> As all searchable claims could be searched without effect justifying an additional fee, this Authority did not invite payment of any additional fee. 3. <input type="checkbox"/> As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.: _____ 4. <input checked="" type="checkbox"/> No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claim Nos.: _____ <p style="text-align: right; margin-top: 10px;">1-20, 24-30</p>
<p>Remark on Prior Art</p> <div style="display: flex; justify-content: space-between;"> <div> <input type="checkbox"/> The additional search fees were accompanied by the applicant's protest. <input type="checkbox"/> No protest accompanied the payment of additional search fees. </div> </div>	

Form PCT/ISA/210 (containing sheet 11) (July 1998)

International Application No. PCT/JP 98/00893

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-28, 24-30

Enhancement of a source coding system comprising an encoder discarding frequency bands, a decoder performing a transposition of frequency bands from a 1st signal to a 2nd signal and a combiner to generate an output signal from 1st and 2nd signal

2. Claims: 21-23, 31, 32

Method and apparatus for computationally efficient transposition by expanding or compressing a time signal

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No. PCT/JP 98/00893

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5127054 A	30-06-1992	NONE	
US 4667340 A	19-05-1987	EP 0124728 A	14-11-1984
		JP 1807228 C	10-12-1993
		JP 5016599 B	04-03-1993
		JP 60035799 A	23-02-1985
US 5068899 A	26-11-1991	CA 1220282 A	07-04-1987

フロントページの続き

- (31) 優先権主張番号 9800268-6
(32) 優先日 平成10年1月30日(1998. 1. 30)
(33) 優先権主張国 スウェーデン (SE)
- (81) 指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, I T, LU, MC, NL, PT, SE), OA(BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), AP(GH, GM, KE, L S, MW, SD, SZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, E E, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, M D, MG, MK, MN, MW, MX, NO, NZ, P L, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, U Z, VN, YU, ZW
- (72) 発明者 ヘン, ラルス, フレドリック
スウェーデン国 エス-168 31 プロム
マ, リタルバサン 14
- (72) 発明者 クヨルリッング, ハンス, マグヌス, クリス
トフェル
スウェーデン国 エス-752 27 ウプサ
ラ, ビンドヘムスガタン19シー